

ON AVERAGE CONTROL GENERATING FAMILIES FOR SINGULARLY PERTURBED OPTIMAL CONTROL PROBLEMS WITH LONG RUN AVERAGE OPTIMALITY CRITERIA

VLADIMIR GAITSGORY*, LUDMILA MANIC[†], AND SERGEY ROSSOMAKHINE[‡]

Abstract. The paper aims at the development of tools for analysis and construction of near optimal solutions of singularly perturbed (SP) optimal controls problems with long run average optimality criteria. The idea that we exploit is to first asymptotically approximate a given problem of optimal control of the SP system by a certain averaged optimal control problem, then reformulate this averaged problem as an infinite-dimensional (ID) linear programming (LP) problem, and then approximate the latter by semi-infinite LP problems. We show that the optimal solution of these semi-infinite LP problems and their duals (that can be found with the help of a modification of an available LP software) allow one to construct near optimal controls of the SP system. We demonstrate the construction with a numerical example.

Key words. Singularly perturbed optimal control problems, Averaging and linear programming, Occupational measures, Numerical solution

AMS subject classifications. 34E15, 34C29, 34A60, 93C70

1. Introduction and preliminaries. Problems of optimal control of singularly perturbed (SP) systems have been studied intensively in both deterministic and stochastic settings (see [2], [8], [17], [18], [23], [26], [28], [29], [33], [41], [52], [55], [56], [58], [61], [63], [68], [67], [65], [74], [75], [77] for a sample of the literature). Originally, the most common approaches to SP control systems, especially in the deterministic case, were related to an approximation of the slow dynamics by the solutions of the systems obtained via equating of the singular perturbations parameter to zero, with further application of the boundary layer method (see [64], [73]) for an asymptotical description of the fast dynamics. This type of approaches were successfully applied to a number of important classes of problems (see, e.g. [17], [29], [55], [56], [58], [63], [65], [74], [77]).

Various averaging type approaches allowing a consideration of more general classes of SP problems, in which the optimal and near optimal controls take the form of rapidly oscillating functions and in which equating of the small parameter to zero does not lead to a right approximation, were studied in [1], [2], [6], [7], [8], [9], [10], [18], [19], [20], [27], [28], [33], [34], [38], [39], [41], [42], [44], [51], [52], [68], [67], [70], [75] (see also references therein). This research lead to a good understanding of what the “true limit” problems, optimal solutions of which approximate optimal solutions of the SP problems, are. However, till recently, no algorithms for finding such approximating solutions (in case fast oscillations may lead to a significant improvement of the performance) have been discussed in the literature, and (to the best of our knowledge) first steps in this direction have been made in the recent publication [48].

The present paper continues the line of research started in [48]. As in [48], our consideration is based on earlier results on averaging of SP control systems obtained in [39], [41], [42], [44], [47] (see also [6], [7], [8], [10], [51], [52], [67]) and on results obtained in [35], [45], [46], [47] that establish the equivalence of optimal control problems to certain infinite dimensional (ID) linear programming (LP) problems (related results on IDLP formulations of optimal control problems in both deterministic and stochastic settings can be found in [5], [15], [16], [21], [31], [37], [50], [53], [57], [60], [69], [71], [72] and [76]). In contrast to [48], where mostly optimal control problems with time discounting criteria were dealt with, this paper is devoted to the consideration of the problems with long run average optimality criteria.

As in [48], we, first, asymptotically approximate a given problem of optimal control of the SP system by a certain averaged optimal control problem, then reformulate this averaged problem as an IDLP problem,

*Department of Mathematics, Macquarie University, Sydney 2109, Australia, vladimir.gaitsgory@mq.edu.au; The work of V. Gaitsgory was supported by the Australian Research Council Discovery Grants DP130104432 and DP120100532

[†]Department of Mathematics, Macquarie University, NSW 2109, Australia, ludmila.manic@mq.edu.au

[‡]Flinders Mathematical Sciences Laboratory, School of Computer Science, Engineering and Mathematics, Flinders University, GPO Box 2100, Adelaide SA 5001, Australia, sergei.rossomakhine@flinders.edu.au; The work of S. Rossomakhine was supported by the Australian Research Council Discovery Grant DP120100532

and then approximate the latter by semi-infinite LP problems. We show that the optimal solution of these semi-infinite LP problems and their duals (that can be found with the help of a modification of an available LP software) allow one to construct near optimal controls of the SP system. Note that, while the approach we exploit is similar to that of [48], the results of this paper are obtained under different assumptions and require a more elaborated argument than those used in [48].

The paper is organized as follows. It consists of eight sections. Section 1 is this introduction. In Section 2, we establish some basic relationships between the SP and the averaged optimal control problems and their IDLP counterparts (Propositions 2.5 and 2.6). In Section 3, the concept of the average control generating (ACG) families for SP problems with long run average criteria is introduced (Definition 3.1) and sufficient and necessary conditions for an ACG family to be optimal are established under the assumption that solutions of the averaged and associated dual problems exist (Proposition 3.5 and Remark 3.6). In Section 4, an approximating averaged semi-infinite LP problem is introduced and it is shown that solutions of the corresponding averaged and associated dual problems exist under natural controllability conditions (Proposition 4.4). In Section 5, it is established that, if certain assumptions are satisfied, then solutions of the approximating averaged and associated dual problems can be used for the construction of near optimal ACG families (Theorem 5.8). In Section 6, we indicate a way how asymptotically near optimal controls of the SP problems with long run time average criteria can be constructed on the basis of near optimal ACG families (Theorem 6.3), the construction being illustrated with a numerical example. In Sections 7 and 8, we give proofs of Theorem 5.8 and Theorem 6.3.

Let us conclude this section with some notations and definitions. Given a compact metric space X , $\mathcal{B}(X)$ will stand for the σ -algebra of its Borel subsets and $\mathcal{P}(X)$ will denote the set of probability measures defined on $\mathcal{B}(X)$. The set $\mathcal{P}(X)$ will always be treated as a compact metric space with a metric ρ , which is consistent with its weak* topology. That is, a sequence $\nu^k \in \mathcal{P}(X)$, $k = 1, 2, \dots$, converges to $\nu \in \mathcal{P}(X)$ in this metric if and only if

$$\lim_{k \rightarrow \infty} \int_X h(x) \nu^k(dx) = \int_X h(x) \nu(dx)$$

for any continuous $h(x) : X \rightarrow \mathbb{R}^1$. Using this metric ρ , one can define the Hausdorff metric ρ_H on the set of closed subsets of $\mathcal{P}(X)$ as follows: $\forall \Gamma_i \subset \mathcal{P}(X)$, $i = 1, 2$,

$$\rho_H(\Gamma_1, \Gamma_2) \stackrel{\text{def}}{=} \max \left\{ \sup_{\nu \in \Gamma_1} \rho(\nu, \Gamma_2), \sup_{\nu \in \Gamma_2} \rho(\nu, \Gamma_1) \right\}, \quad (1.1)$$

where $\rho(\nu, \Gamma_i) \stackrel{\text{def}}{=} \inf_{\nu' \in \Gamma_i} \rho(\nu, \nu')$.

Given a measurable function $x(\cdot) : [0, \infty) \rightarrow X$, the *occupational measure* generated by this function on the interval $[0, S]$ is the probability measure $\nu^{x(\cdot), S} \in \mathcal{P}(X)$ defined by the equation

$$\nu^{x(\cdot), S}(\mathbf{B}) \stackrel{\text{def}}{=} \frac{1}{S} \int_0^S 1_{\mathbf{B}}(x(t)) dt, \quad \forall \mathbf{B} \in \mathcal{B}(X), \quad (1.2)$$

where $1_{\mathbf{B}}(\cdot)$ is the indicator function. The occupational measure generated by this function on the interval $[0, \infty)$ is the probability measure $\nu^{x(\cdot)} \in \mathcal{P}(X)$ defined as the limit (assumed to exist)

$$\nu^{x(\cdot)}(\mathbf{B}) \stackrel{\text{def}}{=} \lim_{S \rightarrow \infty} \frac{1}{S} \int_0^S 1_{\mathbf{B}}(x(t)) dt, \quad \forall \mathbf{B} \in \mathcal{B}(X). \quad (1.3)$$

Note that (1.2) is equivalent to that

$$\int_X h(x) \nu^{x(\cdot), S}(dx) = \frac{1}{S} \int_0^S h(x(t)) dt \quad (1.4)$$

for any $h(\cdot) \in C(X)$, and (1.3) is equivalent to that

$$\int_X h(x) \nu^{x(\cdot)}(dx) = \lim_{S \rightarrow \infty} \frac{1}{S} \int_0^S h(x(t)) dt \quad (1.5)$$

for any $h(\cdot) \in C(X)$.

2. Singularly perturbed and averaged optimal control problems and the related IDLP problems. Consider the SP control system

$$\epsilon y'(t) = f(u(t), y(t), z(t)), \quad (2.1)$$

$$z'(t) = g(u(t), y(t), z(t)), \quad (2.2)$$

where $\epsilon > 0$ is a small parameter; $f(\cdot) : U \times \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g(\cdot) : U \times \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ are continuous vector functions satisfying Lipschitz conditions in z and y ; and where controls $u(\cdot)$ are measurable functions of time satisfying the inclusion

$$u(t) \in U, \quad (2.3)$$

U being a given compact metric space.

Let Y be a given compact subset of \mathbb{R}^m and Z be a given compact subset of \mathbb{R}^n such that the system (2.1)-(2.2) is viable in $Y \times Z$ for any $\epsilon > 0$ small enough (see the definition of viability in [13]).

DEFINITION 2.1. Let $u(\cdot)$ be a control and let $(y_\epsilon(\cdot), z_\epsilon(\cdot))$ be the corresponding solution of the system (2.1)-(2.2). The triplet $(u(\cdot), y_\epsilon(\cdot), z_\epsilon(\cdot))$ will be called *admissible* if

$$(y_\epsilon(t), z_\epsilon(t)) \in Y \times Z \quad \forall t \geq 0. \quad (2.4)$$

In this paper we will be dealing with the optimal control problem

$$\inf_{(u(\cdot), y_\epsilon(\cdot), z_\epsilon(\cdot))} \liminf_{T \rightarrow \infty} \frac{1}{T} \int_0^T G(u(t), y_\epsilon(t), z_\epsilon(t)) dt \stackrel{\text{def}}{=} V^*(\epsilon), \quad (2.5)$$

where $G(\cdot)$ is a continuous function and *inf* is over all admissible triplets of the SP system. Note that the initial conditions are not fixed in (2.1)-(2.2) and they are, in fact, a part of the optimization problem. Note also that, under natural conditions, the optimal value of the problem (2.5) is equal to the optimal value of the periodic optimization problem

$$\inf_{(\mathcal{T}, u(\cdot), y_\epsilon(\cdot), z_\epsilon(\cdot))} \frac{1}{T} \int_0^T G(u(t), y_\epsilon(t), z_\epsilon(t)) dt, \quad (2.6)$$

where *inf* is over the length \mathcal{T} of the time interval and over the admissible triplets that are defined on this interval and that satisfy the periodicity condition: $(y_\epsilon(\mathcal{T}), z_\epsilon(\mathcal{T})) = (y_\epsilon(0), z_\epsilon(0))$. Although the periodic optimization formulation seems to be simpler, a more general statement of the problem in the form (2.5) is more convenient for our consideration.

The SP optimal control problem (2.5) is related to the infinite dimensional linear programming problem

$$\min_{\gamma \in \mathcal{W}(\epsilon)} \left\{ \int_{U \times Y \times Z} G(u, y, z) \gamma(du, dy, dz) \right\} \stackrel{\text{def}}{=} G^*(\epsilon), \quad (2.7)$$

where

$$\begin{aligned} \mathcal{W}(\epsilon) \stackrel{\text{def}}{=} \{ \gamma \in \mathcal{P}(U \times Y \times Z) : \int_{U \times Y \times Z} \nabla(\phi(y)\psi(z))^T \chi_\epsilon(u, y, z) \gamma(du, dy, dz) = 0 \\ \forall \phi(\cdot) \in C^1(\mathbb{R}^m), \quad \forall \psi(\cdot) \in C^1(\mathbb{R}^n) \}, \end{aligned} \quad (2.8)$$

with $\chi_\epsilon(u, y, z)^T \stackrel{\text{def}}{=} (\frac{1}{\epsilon} f(u, y, z)^T, g(u, y, z)^T)$. Namely, the optimal values of these two problems are related by the inequality

$$V^*(\epsilon) \geq G^*(\epsilon) \quad \forall \epsilon > 0, \quad (2.9)$$

and also, under certain conditions (see [35] and [47]),

$$V^*(\epsilon) = G^*(\epsilon) \quad \forall \epsilon > 0. \quad (2.10)$$

Along with the SP system (2.1)-(2.2), let us consider a so-called associate system

$$y'(\tau) = f(u(\tau), y(\tau), z), \quad z = \text{const}. \quad (2.11)$$

Note that the associated system (2.11) looks similar to the “fast” subsystem (2.1) but, in contrast to (2.1), it is evolving in the “stretched” time scale $\tau = \frac{t}{\epsilon}$, with z being a vector of fixed parameters. Everywhere in what follows, it is assumed that the associated system is viable in Y .

DEFINITION 2.2. *A pair $(u(\cdot), y(\cdot))$ will be called admissible for the associated system if (2.11) is satisfied for almost all τ ($u(\cdot)$ being measurable and $y(\cdot)$ being absolutely continuous functions) and if*

$$u(\tau) \in U, \quad y(\tau) \in Y \quad \forall \tau \geq 0. \quad (2.12)$$

Denote by $\mathcal{M}(z, S, y)$ the set of occupational measures generated on the interval $[0, S]$ by the admissible pairs of the associated system that satisfy the initial conditions $y(0) = y$. That is,

$$\mathcal{M}(z, S, y) \stackrel{\text{def}}{=} \bigcup_{(u(\cdot), y(\cdot))} \{\mu^{(u(\cdot), y(\cdot)), S}\} \subset \mathcal{P}(U \times Y),$$

where $\mu^{(u(\cdot), y(\cdot)), S}$ is the occupational measure generated on the interval $[0, S]$ by an admissible pair of the associated system $(u(\cdot), y(\cdot))$ satisfying the initial condition $y(0) = y$ and the union is over such admissible pairs. Also, denote by $\mathcal{M}(z, S)$ the union of $\mathcal{M}(z, S, y)$ over all $y \in Y$,

$$\mathcal{M}(z, S) \stackrel{\text{def}}{=} \bigcup_{y \in Y} \{\mathcal{M}(z, S, y)\}.$$

In [41] it has been established that

$$\limsup_{S \rightarrow \infty} \bar{co} \mathcal{M}(z, S) \subset W(z) \quad (2.13)$$

and that, under mild conditions,

$$\lim_{S \rightarrow \infty} \rho_H(\bar{co} \mathcal{M}(z, S), W(z)) = 0, \quad (2.14)$$

where \bar{co} stands for the closed convex hull of the corresponding set and $W(z) \subset \mathcal{P}(U \times Y)$ is defined by the equation

$$W(z) \stackrel{\text{def}}{=} \{\mu \in \mathcal{P}(U \times Y) : \int_{U \times Y} \nabla \phi(y)^T f(u, y, z) \mu(du, dy) = 0 \quad \forall \phi(\cdot) \in C^1(\mathbb{R}^m)\} \quad (2.15)$$

(see Theorem 2.1(i) in [41]). Also, it has been established that, under some additional conditions (see Theorem 2.1(ii),(iii) and Proposition 4.1 in [41]),

$$\lim_{S \rightarrow \infty} \rho_H(\mathcal{M}(z, S, y), W(z)) = 0 \quad \forall y \in Y, \quad (2.16)$$

with the convergence being uniform with respect to $y \in Y$.

Define the function $\tilde{g}(\mu, z) : \mathcal{P}(U \times Y) \times Z \rightarrow \mathbb{R}^n$ by the equation

$$\tilde{g}(\mu, z) \stackrel{\text{def}}{=} \int_{U \times Y} g(u, y, z) \mu(du, dy) \quad \forall \mu \in \mathcal{P}(U \times Y) \quad (2.17)$$

and consider the system

$$z'(t) = \tilde{g}(\mu(t), z(t)), \quad (2.18)$$

in which the role of controls is played by measure valued functions $\mu(\cdot)$ that satisfy the inclusion

$$\mu(t) \in W(z(t)). \quad (2.19)$$

The system (2.18) will be referred to as *the averaged system*. In what follows, it is assumed that the averaged system is viable in Z .

DEFINITION 2.3. A pair $(\mu(\cdot), z(\cdot))$ will be referred to as admissible for the averaged system if (2.18) and (2.19) are satisfied for almost all t ($\mu(\cdot)$ being measurable and $z(\cdot)$ being absolutely continuous functions) and if

$$z(t) \in Z \quad \forall t \geq 0. \quad (2.20)$$

From Theorem 2.8 of [44] (see also Corollary 3.1 in [41]) it follows that, under the assumption that (2.16) is satisfied (and under other assumptions including the Lipschitz continuity of the multi-valued map $V(z) \stackrel{\text{def}}{=} \cup_{\mu \in W(z)} \{\tilde{g}(\mu, z)\}$), the averaged system approximates the SP dynamics on the infinite time horizon in the sense that the following two statements are valid:

(i) Given an admissible triplet $(u(\cdot), y_\epsilon(\cdot), z_\epsilon(\cdot))$ of the SP system (2.1)-(2.2) that satisfies the initial condition

$$(y_\epsilon(0), z_\epsilon(0)) = (y_0, z_0), \quad (2.21)$$

there exists an admissible pair of the averaged system $(\mu(\cdot), z(\cdot))$ satisfying the initial condition

$$z(0) = z_0 \quad (2.22)$$

such that

$$\sup_{t \in [0, \infty)} \|z_\epsilon(t) - z(t)\| \leq \alpha(\epsilon), \quad \text{where} \quad \lim_{\epsilon \rightarrow 0} \alpha(\epsilon) = 0 \quad (2.23)$$

and, for any Lipschitz continuous functions $h(u, y, z)$,

$$\sup_{T > 0} \left| \frac{1}{T} \int_0^T h(u(t), y_\epsilon(t), z_\epsilon(t)) dt - \frac{1}{T} \int_0^T \tilde{h}(\mu(t), z(t)) dt \right| \leq \alpha_h(\epsilon), \quad \text{where} \quad \lim_{\epsilon \rightarrow 0} \alpha_h(\epsilon) = 0 \quad (2.24)$$

where

$$\tilde{h}(\mu, z) \stackrel{\text{def}}{=} \int_{U \times Y} h(u, y, z) \mu(du, dy) \quad \forall \mu \in \mathcal{P}(U \times Y). \quad (2.25)$$

(ii) Let $(\mu(\cdot), z(\cdot))$ be an admissible pair of the averaged system satisfying the initial condition (2.22). There exists an admissible triplet $(u(\cdot), y_\epsilon(\cdot), z_\epsilon(\cdot))$ of the SP system satisfying the initial condition (2.21) such that the estimates (2.23) and (2.24) are true.

Without going into technical details, let us introduce the following definition.

DEFINITION 2.4. The averaged system will be said to uniformly approximate the SP system if the statements (i) and (ii) are valid, with the estimates (2.23) and (2.24) being uniform with respect to the initial conditions $(y_0, z_0) \in Y \times Z$.

Consider the optimal control problem

$$\inf_{(\mu(\cdot), z(\cdot))} \liminf_{T \rightarrow \infty} \frac{1}{T} \int_0^T \tilde{G}(\mu(t), z(t)) dt \stackrel{\text{def}}{=} \tilde{V}^*, \quad (2.26)$$

where

$$\tilde{G}(\mu, z) \stackrel{\text{def}}{=} \int_{U \times Y} G(u, y, z) \mu(du, dy) \quad (2.27)$$

and where \inf is sought over all admissible pairs of the averaged system (2.18). This will be referred to as *averaged optimal control problem*.

PROPOSITION 2.5. If the averaged system uniformly approximates the SP system, then

$$\lim_{\epsilon \rightarrow 0} V^*(\epsilon) = \tilde{V}^*. \quad (2.28)$$

Proof. The proof follows from the validity of (2.24) (taken with $h(u, y, z) = G(u, y, z)$). \square

The optimal control problem (2.26) is related to the infinite dimensional linear programming problem

$$\tilde{G}^* \stackrel{\text{def}}{=} \min_{p \in \tilde{\mathcal{W}}} \int_F \tilde{G}(\mu, z) p(d\mu, dz) \quad (2.29)$$

where F is the graph of $W(\cdot)$,

$$F \stackrel{\text{def}}{=} \{(\mu, z) : \mu \in W(z), z \in Z\} \subset \mathcal{P}(U \times Y) \times Z, \quad (2.30)$$

and the set $\tilde{\mathcal{W}}$ is defined by the equation

$$\tilde{\mathcal{W}} \stackrel{\text{def}}{=} \{p \in \mathcal{P}(F) : \int_F \nabla \psi(z)^T \tilde{g}(\mu, z) p(d\mu, dz) = 0 \quad \forall \psi(\cdot) \in C^1(\mathbb{R}^n)\}. \quad (2.31)$$

For convenience, this will be referred to as *averaged IDLP problem*. The relationships between the problems (2.26) and (2.29) include, in particular, the inequality between the optimal values

$$\tilde{V}^* \geq \tilde{G}^*, \quad (2.32)$$

which, under certain conditions (see [46]), takes the form of the equality

$$\tilde{V}^* = \tilde{G}^*. \quad (2.33)$$

PROPOSITION 2.6. *The following relationships are valid*

$$\liminf_{\epsilon \rightarrow 0} V^*(\epsilon) \geq \liminf_{\epsilon \rightarrow 0} G^*(\epsilon) \geq \tilde{G}^*. \quad (2.34)$$

If the averaged system uniformly approximates the SP system and if (2.33) is valid, then

$$\lim_{\epsilon \rightarrow 0} V^*(\epsilon) = \lim_{\epsilon \rightarrow 0} G^*(\epsilon) = \tilde{G}^*. \quad (2.35)$$

Note that the first inequality in (2.34) follows from (2.9). The validity of (2.35) follows from (2.28), (2.33) and the second inequality in (2.34). The latter is proved on the basis of the two lemmas stated below.

Having in mind the fact that an arbitrary probability measure $\gamma \in \mathcal{P}(U \times Y \times Z)$ can be “disintegrated” as follows

$$\gamma(du, dy, dz) = \mu(du, dy|z) \nu(dz), \quad (2.36)$$

let us define the set of probability measures $\mathcal{W} \subset \mathcal{P}(U \times Y \times Z)$ by the equation

$$\begin{aligned} \mathcal{W} = \{ & \gamma = \mu(du, dy|z) \nu(dz) : \mu(\cdot|z) \in W(z) \text{ for } \nu - \text{almost all } z \in Z, \\ & \int_Z \nabla \psi(z)^T \tilde{g}(\mu(\cdot|z), z) \nu(dz) = 0 \quad \forall \psi(\cdot) \in C^1(\mathbb{R}^n)\}, \end{aligned} \quad (2.37)$$

where

$$\tilde{g}(\mu(\cdot|z), z) = \int_{U \times Y} g(u, y, z) \mu(du, dy|z). \quad (2.38)$$

Note that the disintegration (2.36) is understood in the sense that, for any continuous $h(u, y, z)$,

$$\int_{U \times Y \times Z} h(u, y, z) \gamma(du, dy, dz) = \int_Z \left(\int_{U \times Y} h(u, y, z) \mu(du, dy|z) \right) \nu(dz),$$

where $\nu(dz) \stackrel{\text{def}}{=} \gamma(U \times Y, dz)$, $\mu(du, dy|z)$ is a probability measure on $U \times Y$ such that the integral $\int_{U \times Y} h(u, y, z) \mu(du, dy|z)$ is Borel measurable on Z .

LEMMA 2.7. *The following relationship is valid:*

$$\limsup_{\epsilon \rightarrow 0} \mathcal{W}(\epsilon) \subset \mathcal{W}. \quad (2.39)$$

Proof. The proof follows from Propositions 4.1 and 4.2 in [48]. \square

Define the map $\Phi(\cdot) : \mathcal{P}(F) \rightarrow \mathcal{P}(U \times Y \times Z)$ as follows. For any $p \in \mathcal{P}(F)$, let $\Phi(p) \in \mathcal{P}(U \times Y \times Z)$ be such that

$$\int_{U \times Y \times Z} h(u, y, z) \Phi(p)(du, dy, dz) = \int_F \tilde{h}(\mu, z) p(d\mu, dz) \quad \forall h(\cdot) \in C(U \times Y \times Z), \quad (2.40)$$

where $\tilde{h}(\mu, z) = \int_{U \times Y} h(u, y, z) \mu(du, dy)$ (this definition is legitimate since the right-hand side of the above expression defines a linear continuous functional on $C(U \times Y \times Z)$, the latter being associated with an element of $\mathcal{P}(U \times Y \times Z)$ that makes the equality (2.40) valid). Note that the map $\Phi(\cdot) : \mathcal{P}(F) \rightarrow \mathcal{P}(U \times Y \times Z)$ is linear and it is continuous in the sense that

$$\lim_{p_l \rightarrow p} \Phi(p_l) = \Phi(p), \quad (2.41)$$

with p_l converging to p in the weak* topology of $\mathcal{P}(F)$ and $\Phi(p_l)$ converging to $\Phi(p)$ in the weak* topology of $\mathcal{P}(U \times Y \times Z)$ (see Lemma 4.3 in [44]).

LEMMA 2.8. *The following equality is true:*

$$\mathcal{W} = \Phi(\tilde{\mathcal{W}}). \quad (2.42)$$

Proof. The proof follows from Proposition 5.6 in [48]. \square

Proof of Proposition 2.6. As was mentioned above, to prove the proposition, it is sufficient to establish the validity of the second inequality in (2.34). Note that, by (2.39),

$$\liminf_{\epsilon \rightarrow 0} G^*(\epsilon) \geq \min_{\gamma \in \mathcal{W}} \left\{ \int_{U \times Y \times Z} G(u, y, z) \gamma(du, dy, dz) \right\}. \quad (2.43)$$

Also, by (2.42),

$$\min_{\gamma \in \mathcal{W}} \left\{ \int_{U \times Y \times Z} G(u, y, z) \gamma(du, dy, dz) \right\} = \min_{p \in \tilde{\mathcal{W}}} \left\{ \int_{U \times Y \times Z} G(u, y, z) \Phi(p)(du, dy, dz) \right\} = \tilde{G}^* \quad (2.44)$$

(the last equality being due to (2.40)). By comparing (2.43) and (2.44), one obtains the second inequality in (2.34). \square

3. Average control generating (ACG) families. For any $z \in Z$, let $(u_z(\cdot), y_z(\cdot))$ be an admissible pair of the associated system (2.11) and $\mu(du, dy|z)$ be the occupational measure generated by this pair on $[0, \infty)$ (see (1.5)), with the integral $\int_{U \times Y} h(u, y, z) \mu(du, dy|z)$ being a measurable function of z and

$$\left| \frac{1}{S} \int_0^S h(u_z(\tau), y_z(\tau), z) d\tau - \int_{U \times Y} h(u, y, z) \mu(du, dy|z) \right| \leq \phi_h(S) \quad \forall z \in Z, \quad \lim_{S \rightarrow \infty} \phi_h(S) = 0 \quad (3.1)$$

for any continuous function $h(u, y, z) : U \times Y \times Z \rightarrow \mathbb{R}^1$. Note that, due to (2.13),

$$\mu(du, dy|z) \in W(z) \quad \forall z \in Z. \quad (3.2)$$

DEFINITION 3.1. *The family $(u_z(\cdot), y_z(\cdot))$ will be called average control generating (ACG) if the system*

$$z'(t) = \tilde{g}_\mu(z(t)), \quad z(0) = z_0, \quad (3.3)$$

where

$$\tilde{g}_\mu(z) \stackrel{\text{def}}{=} \tilde{g}(\mu(\cdot|z), z) = \int_{U \times Y} g(u, y, z) \mu(du, dy|z), \quad (3.4)$$

has a unique solution $z(t) \in Z \forall t \in [0, \infty)$ and, for any continuous function $\tilde{h}(\mu, z) : F \rightarrow \mathbb{R}^1$, there exists a limit

$$\lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{h}(\mu(t), z(t)) dt, \quad (3.5)$$

where $\mu(t) \stackrel{\text{def}}{=} \mu(du, dy|z(t))$.

Note that, according to this definition, if $(u_z(\cdot), y_z(\cdot))$ is an ACG family, with $\mu(du, dy|z)$ being the family of occupational measures generated by this family, and if $z(\cdot)$ is the corresponding solution of (3.3), then the pair $(\mu(\cdot), z(\cdot))$, where $\mu(t) \stackrel{\text{def}}{=} \mu(du, dy|z(t))$, is an admissible pair of the averaged system (for convenience, this admissible pair will also be referred to as one generated by the ACG family). From the fact that the limit (3.5) exists for any continuous $\tilde{h}(\mu, z)$ it follows that the pair $(\mu(\cdot), z(\cdot))$ generates the occupational measure $p \in \mathcal{P}(F)$ defined by the equation

$$\int_F \tilde{h}(\mu, z) p(d\mu, dz) = \lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{h}(\mu(t), z(t)) dt \quad \forall \tilde{h}(\mu, z) \in C(F). \quad (3.6)$$

Also, the state trajectory $z(\cdot)$ generates the occupation measure $\nu \in \mathcal{P}(Z)$ defined by the equation

$$\int_Z h(z) \nu(dz) = \lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} h(z(t)) dt \quad \forall h(z) \in C(Z). \quad (3.7)$$

PROPOSITION 3.2. *Let $(u_z(\cdot), y_z(\cdot))$ be an ACG family and let $\mu(du, dy|z)$ and $(\mu(\cdot), z(\cdot))$ be, respectively, the family of occupational measures and the admissible pair of the averaged system generated by this family. Let p be the occupational measure generated by $(\mu(\cdot), z(\cdot))$ and ν be the occupational measure generated by $z(\cdot)$ (in accordance with (3.6) and (3.7) respectively). Then*

$$p \in \tilde{W} \quad (3.8)$$

and

$$\Phi(p) = \mu(du, dy|z) \nu(dz), \quad (3.9)$$

where $\Phi(\cdot)$ is defined by (2.40).

Proof. For an arbitrary $\psi(\cdot) \in C^1(\mathbb{R}^n)$,

$$\lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \nabla \psi(z(t))^T \tilde{g}(\mu(t), z(t)) dt = \lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} (\psi(z(\mathcal{T})) - \psi(z(0))) = 0.$$

Hence, by (3.6),

$$\int_F \nabla \psi(z)^T \tilde{g}(\mu, z) p(d\mu, dz) = 0 \quad \psi(\cdot) \in C^1(\mathbb{R}^n).$$

The latter implies (3.8). To prove (3.9), note that, for an arbitrary continuous $h(u, y, z)$ and $\tilde{h}(\mu, z)$ defined in accordance with (2.25), one can write down

$$\begin{aligned} \int_F \tilde{h}(\mu, z) p(d\mu, dz) &= \lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{h}(\mu(t), z(t)) dt \\ &= \lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \left(\int_{U \times Y} h(u, y, z) \mu(du, dy|z(t)) \right) dt = \int_Z \left(\int_{U \times Y} h(u, y, z) \mu(du, dy|z) \right) \nu(dz). \end{aligned}$$

By the definition of $\Phi(\cdot)$ (see (2.40)), the latter implies (3.9). \square

DEFINITION 3.3. *An ACG family $(u_z(\cdot), y_z(\cdot))$ will be called optimal if the admissible pair $(\mu(\cdot), z(\cdot))$ generated by this family is optimal in the averaged problem (2.26). That is,*

$$\lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{G}(\mu(t), z(t)) dt = \tilde{V}^*. \quad (3.10)$$

An ACG family $(u_z(\cdot), y_z(\cdot))$ will be called α -near optimal ($\alpha > 0$) if

$$\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau \tilde{G}(\mu(t), z(t)) dt \leq \tilde{V}^* + \alpha. \quad (3.11)$$

COROLLARY 3.4. Let the equality (2.33) be valid. An ACG family $(u_z(\cdot), y_z(\cdot))$ generating the admissible pair $(\mu(\cdot), (z(\cdot)))$ will be optimal if and only if the occupational measure generated by this pair (according to (3.6)) is an optimal solution of the averaged IDLP problem (2.29).

Let $\tilde{H}(p, z)$ be the Hamiltonian corresponding to the averaged optimal control problem (2.26)

$$\tilde{H}(p, z) \stackrel{\text{def}}{=} \min_{\mu \in W(z)} \{ \tilde{G}(\mu, z) + p^T \tilde{g}(\mu, z) \}, \quad (3.12)$$

where $\tilde{g}(\mu, z)$ and $\tilde{G}(\mu, z)$ are defined by (2.17) and (2.27). Consider the problem

$$\sup_{\zeta(\cdot) \in C^1} \{ \theta : \theta \leq \tilde{H}(\nabla \zeta(z), z) \ \forall z \in Z \} = \tilde{G}^*, \quad (3.13)$$

where *sup* is sought over all continuously differentiable functions $\zeta(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^1$. Note that the optimal value of the problem (3.13) is equal to the optimal value of the averaged IDLP problem (2.29). The former is in fact dual with respect to the later, the equality of the optimal values being one of the duality relationships between the two (see Theorem 4.1 in [35]). For brevity, (3.13) will be referred to as just *averaged dual problem*. Note that the averaged dual problem can be equivalently rewritten in the form

$$\sup_{\zeta(\cdot) \in C^1(\mathbb{R}^n)} \{ \theta : \theta \leq \tilde{G}(\mu, z) + \nabla \zeta(z)^T \tilde{g}(\mu, z) \ \forall (\mu, z) \in F \} = \tilde{G}^*, \quad (3.14)$$

where F is the graph of $W(\cdot)$ (see (2.30)). A function $\zeta^*(\cdot) \in C^1$ will be called a solution of the averaged dual problem if

$$\tilde{G}^* \leq \tilde{H}(\nabla \zeta^*(z), z) \ \forall z \in Z, \quad (3.15)$$

or, equivalently, if

$$\tilde{G}^* \leq \tilde{G}(\mu, z) + \nabla \zeta^*(z)^T \tilde{g}(\mu, z) \ \forall (\mu, z) \in F. \quad (3.16)$$

Note that, if $\zeta^*(\cdot) \in C^1$ satisfies (3.15), then $\zeta^*(\cdot) + \text{const}$ satisfies (3.15) as well.

Assume that a solution of the averaged dual problem (that is, a functions $\zeta^*(\cdot)$ satisfying (3.15)) exists and consider the problem in the right hand side of (3.12) with $p = \nabla \zeta^*(z)$ rewriting it in the form

$$\min_{\mu \in W(z)} \left\{ \int_{U \times Y} [G(u, y, z) + \nabla \zeta^*(z)^T g(u, y, z)] \mu(du, dz) \right\} = \tilde{H}(\nabla \zeta^*(z), z). \quad (3.17)$$

The latter is an IDLP problem, with the dual of it having the form

$$\begin{aligned} \sup_{\eta(\cdot) \in C^1(\mathbb{R}^m)} \{ \theta : \theta \leq G(u, y, z) + \nabla \zeta^*(z)^T g(u, y, z) + \nabla \eta(y)^T f(u, y, z) \ \forall (u, y) \in U \times Y \} \\ = \tilde{H}(\nabla \zeta^*(z), z), \end{aligned} \quad (3.18)$$

where *sup* is sought over all continuously differentiable functions $\eta(\cdot) : \mathbb{R}^m \rightarrow \mathbb{R}^1$. The optimal values of the problems (3.17) and (3.18) are equal, this being one of the duality relationships between these two problems (see Theorem 4.1 in [35]). The problem (3.18) will be referred to as *associated dual problem*. A function $\eta_z^*(\cdot) \in C^1(\mathbb{R}^m)$ will be called a solution of the problem (3.18) if

$$\tilde{H}(\nabla \zeta^*(z), z) \leq G(u, y, z) + \nabla \zeta^*(z)^T g(u, y, z) + \nabla \eta_z^*(y)^T f(u, y, z) \ \forall (u, y) \in U \times Y. \quad (3.19)$$

Note that from (3.15) and from (3.19) it follows that

$$G(u, y, z) + \nabla \zeta^*(z)^T g(u, y, z) + \nabla \eta_z^*(y)^T f(u, y, z) \geq \tilde{G}^* \ \forall (u, y, z) \in U \times Y \times Z. \quad (3.20)$$

The following result gives sufficient and also (under additional periodicity assumptions) necessary conditions for an ACG family $(u_z(\cdot), y_z(\cdot))$ to be optimal and for the equality (2.33) to be valid.

PROPOSITION 3.5. *Let a solution $\zeta^*(z)$ of the averaged dual problem exist and a solution $\eta_z^*(y)$ of the associated dual problem exist for any $z \in Z$. Then an ACG family $(u_z(\cdot), y_z(\cdot))$ generating the admissible pair of the averaged system $(\mu(\cdot), z(\cdot))$ is optimal and the equality (2.33) is valid if*

$$\begin{aligned} & G(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) + \nabla \zeta^*(z(t))^T g(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) \\ & + \nabla \eta_{z(t)}^*(y_{z(t)}(\tau))^T f(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) = \tilde{G}^* \quad \forall \tau \in P_t, \quad \forall t \in A, \end{aligned} \quad (3.21)$$

for some $P_t \subset \mathbb{R}^1$ and $A \subset \mathbb{R}^1$ such that

$$\text{meas}\{\mathbb{R}^1 \setminus P_t\} = 0 \quad \forall t \in A \quad \text{and} \quad \text{meas}\{\mathbb{R}^1 \setminus A\} = 0. \quad (3.22)$$

Under the additional assumptions that an ACG family $(u_z(\cdot), y_z(\cdot))$ is periodic, that is,

$$(u_z(\tau + T_z), y_z(\tau + T_z)) = (u_z(\tau), y_z(\tau)) \quad \forall \tau \geq 0 \quad (3.23)$$

for some $T_z > 0$ and that the admissible pair of the averaged system $(\mu(\cdot), z(\cdot))$ generated by this family is periodic as well, that is,

$$(\mu(t + \tilde{T}), z(t + \tilde{T})) = (\mu(t), z(t)) \quad \forall t \geq 0 \quad (3.24)$$

for some $\tilde{T} > 0$, the fulfillment of (3.21) is also necessary for $(u_z(\cdot), y_z(\cdot))$ to be optimal and for the equality (2.33) to be valid.

Proof. Assume (3.21) is true. Then

$$\begin{aligned} & \lim_{S \rightarrow \infty} \frac{1}{S} \int_0^S [G(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) + \nabla \zeta^*(z(t))^T g(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) \\ & + \nabla \eta_{z(t)}^*(y_{z(t)}(\tau))^T f(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t))] d\tau \\ & = \tilde{G}(\mu(t), z(t)) + \nabla \zeta^*(z(t))^T \tilde{g}(\mu(t), z(t)) = \tilde{G}^* \quad \forall t \in A, \end{aligned} \quad (3.25)$$

where it has been taken into account that

$$\lim_{S \rightarrow \infty} \frac{1}{S} \int_0^S \nabla \eta_{z(t)}^*(y_{z(t)}(\tau))^T f(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) d\tau = \lim_{S \rightarrow \infty} \frac{1}{S} [\eta_{z(t)}^*(y_{z(t)}(S)) - \eta_{z(t)}^*(y_{z(t)}(0))] = 0.$$

Since

$$\lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \nabla \zeta^*(z(t))^T \tilde{g}(\mu(t), z(t)) dt = \lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} (\zeta^*(z(\mathcal{T})) - \zeta^*(z(0))) = 0,$$

from (3.25) it follows that

$$\lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{G}(\mu(t), z(t)) dt = \tilde{G}^*.$$

By (2.32), the latter implies that $(u_z(\cdot), y_z(\cdot))$ is optimal and that the equality (2.33) is valid.

Let us now prove (assuming that (3.23) and (3.24) are true) that the fulfillment of (3.21) is necessary for an ACG family $(u_z(\cdot), y_z(\cdot))$ to be optimal and for the equality (2.33) to be valid. In fact, let an ACG $(u_z(\cdot), y_z(\cdot))$ be optimal and let (2.33) be true. Then

$$\frac{1}{\tilde{T}} \int_0^{\tilde{T}} \tilde{G}(\mu(t), z(t)) dt = \tilde{G}^*.$$

Since (by (3.24))

$$\int_0^{\tilde{T}} \nabla \zeta^*(z(t))^T \tilde{g}(\mu(t), z(t)) dt = \zeta^*(z(\tilde{T})) - \zeta^*(z(0)) = 0,$$

it follows that

$$\frac{1}{\tilde{T}} \int_0^{\tilde{T}} [\tilde{G}(\mu(t), z(t)) + \nabla \zeta^*(z(t))^T \tilde{g}(\mu(t), z(t)) - \tilde{G}^*] dt = 0$$

and, hence, by (3.16),

$$\tilde{G}(\mu(t), z(t)) + \nabla \zeta^*(z(t))^T \tilde{g}(\mu(t), z(t)) = \tilde{G}^* \quad (3.26)$$

for almost all $t \in [0, \tilde{T}]$. Note that (due to the periodicity condition (3.24)) the equality above is also valid for almost all $t \in [0, \infty)$.

Let the set A ($meas\{\mathbb{R}^1 \setminus A\} = 0$) be such that the equality (3.26) is valid and let $t \in A$. Due to the periodicity condition (3.23), to prove the required statement it is sufficient to show that the equality (3.21) is satisfied for almost all $\tau \in [0, T_{z(t)}]$. Assume it is not the case and there exists a set $Q_t \subset [0, T_{z(t)}]$, with $meas\{Q_t\} > 0$, on which (3.21) is not satisfied, the latter implying (due to (3.20)) that

$$\begin{aligned} & G(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) + \nabla \zeta^*(z(t))^T g(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) \\ & + \nabla \eta_{z(t)}^*(y_{z(t)}(\tau))^T f(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) > \tilde{G}^* \quad \forall \tau \in Q_t. \end{aligned}$$

From the above inequality and from (3.20) it follows that

$$\begin{aligned} & \frac{1}{T_{z(t)}} \int_0^{T_{z(t)}} [G(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) + \nabla \zeta^*(z(t))^T g(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) \\ & + \nabla \eta_{z(t)}^*(y_{z(t)}(\tau))^T f(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) d\tau > \tilde{G}^*. \end{aligned} \quad (3.27)$$

By (3.23),

$$\int_0^{T_{z(t)}} \nabla \eta_{z(t)}^*(y_{z(t)}(\tau))^T f(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) d\tau = \eta_{z(t)}^*(y_{z(t)}(T_{z(t)})) - \eta_{z(t)}^*(y_{z(t)}(0)) = 0. \quad (3.28)$$

Hence, from (3.27) it follows that

$$\frac{1}{T_{z(t)}} \int_0^{T_{z(t)}} [G(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) + \nabla \zeta^*(z(t))^T g(u_{z(t)}(\tau), y_{z(t)}(\tau), z(t))] d\tau > \tilde{G}^*,$$

which is equivalent to

$$\tilde{G}(\mu(t), z(t)) + \nabla \zeta^*(z(t))^T \tilde{g}(\mu(t), z(t)) > \tilde{G}^*.$$

This contradicts to the fact that t was chosen to belong to the set A on which (3.26) is satisfied. This completes the proof of the proposition. \square

REMARK 3.6. *Note that, due to (3.20), the validity of (3.21) implies the validity of the inclusion*

$$\begin{aligned} & (u_{z(t)}(\tau), y_{z(t)}(\tau), z(t)) \in \text{Argmin}_{(u,y,z) \in U \times Y \times Z} \{G(u, y, z) + \nabla \zeta^*(z)^T g(u, y, z) \\ & + \nabla \eta_z^*(y)^T f(u, y, z)\} \quad \forall \tau \in P_t, \quad \forall t \in A \end{aligned} \quad (3.29)$$

which, in turn, implies

$$\begin{aligned} & u_{z(t)}(\tau) \in \text{Argmin}_{u \in U} \{G(u, y_{z(t)}(\tau), z(t)) + \nabla \zeta^*(z(t))^T g(u, y_{z(t)}(\tau), z(t)) \\ & + \nabla \eta_{z(t)}^*(y_{z(t)}(\tau))^T f(u, y_{z(t)}(\tau), z(t))\} \quad \forall \tau \in P_t, \quad \forall t \in A. \end{aligned} \quad (3.30)$$

That is, if the equality (2.33) is valid, then for an ACG family $(u_z(\cdot), y_z(\cdot))$ satisfying the periodicity conditions (3.23) and (3.24) to be optimal, it is necessary that the inclusion (3.30) is satisfied.

4. Approximating averaged semi-infinite LP problem; conditions for the existence of solutions of the approximating averaged/associated dual problems. Let $\psi_i(\cdot) \in C^1(\mathbb{R}^n)$, $i = 1, 2, \dots$, be a sequence of functions such that any $\zeta(\cdot) \in C^1(\mathbb{R}^n)$ and its gradient are simultaneously approximated by a linear combination of $\psi_i(\cdot)$ and their gradients. Also, let $\phi_i(\cdot) \in C^1(\mathbb{R}^m)$, $i = 1, 2, \dots$, be a sequence of functions such that any $\eta(\cdot) \in C^1(\mathbb{R}^m)$ and its gradient are simultaneously approximated by a linear combination of $\phi_i(\cdot)$ and their gradients. Examples of such sequences are monomials $z_1^{i_1} \dots z_n^{i_n}$, $i_1, \dots, i_n = 0, 1, \dots$ and, respectively, $y_1^{i_1} \dots y_m^{i_m}$, $i_1, \dots, i_m = 0, 1, \dots$, with z_k ($k = 1, \dots, n$), and y_l ($l = 1, \dots, m$) standing for the components of z and y (see, e.g., [62]).

Let us introduce the following notations:

$$W_M(z) \stackrel{\text{def}}{=} \{\mu \in \mathcal{P}(U \times Y) : \int_{U \times Y} \nabla \phi_i(y)^T f(u, y, z) \mu(du, dy) = 0, \quad i = 1, \dots, M\}, \quad (4.1)$$

$$F_M \stackrel{\text{def}}{=} \{(\mu, z) : \mu \in W_M(z), \quad z \in Z\} \subset \mathcal{P}(U \times Y) \times Z, \quad (4.2)$$

$$\tilde{\mathcal{W}}^{N,M} \stackrel{\text{def}}{=} \{p \in \mathcal{P}(F_M) : \int_{F_M} \nabla \psi_i(z)^T \tilde{g}(\mu, z) p(d\mu, dz) = 0, \quad i = 1, \dots, N\}, \quad (4.3)$$

(compare with (2.15), (2.30) and (2.31), respectively) and let us consider the following semi-infinite LP problem

$$\min_{p \in \tilde{\mathcal{W}}^{N,M}} \int_{F_M} \tilde{G}(\mu, z) p(d\mu, dz) \stackrel{\text{def}}{=} \tilde{G}^{N,M} \quad (4.4)$$

(compare with (2.29)). This problem will be referred to as (N, M) -approximating averaged problem.

It is obvious that

$$W_1(z) \supset W_2(z) \supset \dots \supset W_M(z) \supset \dots \supset W(z) \quad \Rightarrow \quad F_1 \supset F_2 \supset \dots \supset F_M \supset \dots \supset F. \quad (4.5)$$

Defining the set $\tilde{\mathcal{W}}^N$ by the equation

$$\tilde{\mathcal{W}}^N \stackrel{\text{def}}{=} \{p \in \mathcal{P}(F) : \int_F \nabla \psi_i(z)^T \tilde{g}(\mu, z) p(d\mu, dz) = 0, \quad i = 1, \dots, N\}, \quad (4.6)$$

one can also see that

$$\tilde{\mathcal{W}}^{N,M} \supset \tilde{\mathcal{W}}^N \supset \tilde{\mathcal{W}} \quad \forall N, M = 1, 2, \dots \quad (4.7)$$

(with $\mathcal{W}^{N,M}$, $\tilde{\mathcal{W}}^N$ and $\tilde{\mathcal{W}}$ being considered as subsets of $\mathcal{P}(\mathcal{P}(U \times Y) \times Z)$), the latter implying, in particular, that

$$\tilde{G}^{N,M} \leq \tilde{G}^* \quad \forall N, M = 1, 2, \dots \quad (4.8)$$

It can be readily verified that (see, e.g., the proof of Proposition 7 in [47]) that

$$\lim_{M \rightarrow \infty} W_M(z) = W(z), \quad \lim_{M \rightarrow \infty} F_M = F, \quad (4.9)$$

where, in the first case, the convergence is in the Hausdorff metric generated by the weak convergence in $\mathcal{P}(U \times Y)$ and, in the second, it is in the Hausdorff metric generated by the weak* convergence in $\mathcal{P}(U \times Y)$ and the convergence in Z .

PROPOSITION 4.1. *The following relationships are valid:*

$$\lim_{M \rightarrow \infty} \tilde{\mathcal{W}}^{N,M} = \tilde{\mathcal{W}}^N, \quad \lim_{N \rightarrow \infty} \tilde{\mathcal{W}}^N = \tilde{\mathcal{W}}, \quad (4.10)$$

where the convergence in both cases is in Hausdorff metric generated by the weak* convergence in $\mathcal{P}(\mathcal{P}(U \times Y) \times Z)$. Also,

$$\lim_{N \rightarrow \infty} \lim_{M \rightarrow \infty} \tilde{G}^{N,M} = \tilde{G}^*. \quad (4.11)$$

If the optimal solution p^* of the averaged IDLP problem (2.29) is unique, then, for an arbitrary optimal solution $p^{N,M}$ of the (N, M) -approximating problem (4.4),

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \rho(p^{N,M}, p^*) = 0. \quad (4.12)$$

Proof. The proof is similar to that of Proposition 3.5 in [48]. \square

Define the finite dimensional space $\mathcal{Q}_N \subset C^1(\mathbb{R}^n)$ by the equation

$$\mathcal{Q}_N \stackrel{\text{def}}{=} \{ \zeta(\cdot) \in C^1(\mathbb{R}^n) : \zeta(z) = \sum_{i=1}^N \lambda_i \psi_i(z), \quad \lambda = (\lambda_i) \in \mathbb{R}^N \} \quad (4.13)$$

and consider the following problem

$$\sup_{\zeta(\cdot) \in \mathcal{Q}_N} \{ \theta : \theta \leq \tilde{G}(\mu, z) + \nabla \zeta(z)^T \tilde{g}(\mu, z) \quad \forall (\mu, z) \in F_M \} = \tilde{G}^{N,M}. \quad (4.14)$$

This problem is dual with respect to the problem (4.4), the equality of the optimal values of these two problems being a part of the duality relationships. Note that the problem (4.14) looks similar to the averaged dual problem (3.14). However, in contrast to the latter, in (4.14), \sup is sought over the finite dimensional subspace \mathcal{Q}_N of $C^1(\mathbb{R}^n)$ and F_M is used instead of F . The problem (4.14) will be referred to as (N, M) -approximating averaged dual problem. A function $\zeta^{N,M}(\cdot) \in \mathcal{Q}_N$,

$$\zeta^{N,M}(z) = \sum_{i=1}^N \lambda_i^{N,M} \psi_i(z), \quad (4.15)$$

will be called a solution of the (N, M) -approximating averaged dual problem if

$$\tilde{G}^{N,M} \leq \tilde{G}(\mu, z) + \nabla \zeta^{N,M}(z)^T \tilde{g}(\mu, z) \quad \forall (\mu, z) \in F_M. \quad (4.16)$$

Define the finite dimensional space $\mathcal{V}_M \subset C^1(\mathbb{R}^m)$ by the equation

$$\mathcal{V}_M \stackrel{\text{def}}{=} \{ \eta(\cdot) \in C^1(\mathbb{R}^m) : \eta(y) = \sum_{i=1}^M \omega_i \phi_i(y), \quad \omega = (\omega_i) \in \mathbb{R}^M \} \quad (4.17)$$

and, assuming that a solution $\zeta^{N,M}(z)$ of the (N, M) -approximating averaged dual problem exists, consider the following problem

$$\sup_{\eta(\cdot) \in \mathcal{V}_M} \{ \theta : \theta \leq G(u, y, z) + \nabla \zeta^{N,M}(z)^T g(u, y, z) + \nabla \eta(y)^T f(u, y, z) \quad \forall (u, y) \in U \times Y \} \stackrel{\text{def}}{=} \sigma^{N,M}(z). \quad (4.18)$$

While the problem (4.18) looks similar to the associated dual problem (3.18), it differs from the latter, firstly, by that \sup is sought over the finite dimensional subspace \mathcal{V}_M of $C^1(\mathbb{R}^m)$ and, secondly, by that a solution $\zeta^{N,M}(z)$ of (4.14) is used instead of a solution $\zeta^*(z)$ of (3.13) (the latter may not exist). The problem (4.18) will be referred to as (N, M) -approximating associated dual problem. It can be shown that it is, indeed, dual with respect to the semi-infinite LP problem

$$\min_{\mu \in W_M(z)} \left\{ \int_{U \times Y} [G(u, y, z) + \nabla \zeta^{N,M}(z)^T g(u, y, z)] \mu(du, dy) \right\} = \sigma^{N,M}(z), \quad (4.19)$$

the duality relationships including the equality of the optimal values (see Theorem 5.2(ii) in [35]). A function $\eta_z^{N,M}(\cdot) \in \mathcal{V}_M$,

$$\eta_z^{N,M}(y) = \sum_{i=1}^M \omega_{z,i}^{N,M} \phi_i(y), \quad (4.20)$$

will be called a solution of the (N, M) -approximating associated dual problem if

$$\sigma^{N,M}(z) \leq G(u, y, z) + \nabla \zeta^{N,M}(z)^T g(u, y, z) + \nabla \eta_z^{N,M}(y)^T f(u, y, z) \quad \forall (u, y) \in U \times Y. \quad (4.21)$$

Let us now introduce two controllability type assumptions under which solutions of the (N, M) -approximating averaged and associated dual problems exists.

ASSUMPTION 4.2. *There exists a set $Z^0 \subset Z$, the closure of which has a nonempty interior,*

$$\text{int}(\text{cl}Z^0) \neq \emptyset,$$

such that any two points in Z^0 can be connected by an admissible trajectory of the averaged system (that is, for any $z', z'' \in Z^0$, there exists an admissible pair $(\mu(\cdot), z(\cdot))$ of the averaged system defined on some interval $[0, T]$ such that $z(0) = z'$ and $z(T) = z''$).

ASSUMPTION 4.3. *There exists a set $Y^0(z) \subset Y$, the closure of which has a nonempty interior,*

$$\text{int}(\text{cl}Y^0(z)) \neq \emptyset,$$

such that any two points in $Y^0(z)$ can be connected by an admissible trajectory of the associated system (that is, for any $y', y'' \in Y^0(z)$, there exists an admissible pair $(u(\cdot), y(\cdot))$ of the associated system defined on some interval $[0, S]$ such that $y(0) = y'$ and $y(S) = y''$).

Assume also that, for any $N = 1, 2, \dots$, and $M = 1, 2, \dots$, the gradients $\nabla\psi_i(z)$, $i = 1, 2, \dots, N$, and $\nabla\phi_i(y)$, $i = 1, 2, \dots, M$, are linearly independent on any open subset of \mathbb{R}^N and, respectively, \mathbb{R}^M . That is, if Q is an open subset of \mathbb{R}^N , then the equality

$$\sum_{i=1}^N v_i \nabla\psi_i(z) = 0 \quad \forall z \in Q$$

is valid only if $v_i = 0$, $i = 1, \dots, N$, and, similarly, if D is an open subset of \mathbb{R}^M , then the equality

$$\sum_{i=1}^M w_i \nabla\phi_i(y) = 0 \quad \forall y \in D$$

is valid only if $w_i = 0$, $i = 1, \dots, M$.

PROPOSITION 4.4.

(i) *If Assumption 4.2 is satisfied, then a solution of the (N, M) -approximating averaged dual problem exists for any N and M .*

(ii) *If Assumption 4.3 is satisfied for any $z \in Z$, then a solution of the (N, M) -approximating associated dual problem exists for any N and M , and for any $z \in Z$.*

The proof of the propositions is based on Lemma 3.10 of [48] reproduced below

LEMMA 4.5. *Let X be a compact metric space and let $\Psi_i(\cdot) : X \rightarrow \mathbb{R}^1$, $i = 0, 1, \dots, K$, be continuous functional on X . Let*

$$\sigma^* \stackrel{\text{def}}{=} \sup_{\{\lambda_i\}} \{ \theta : \theta \leq \Psi_0(x) + \sum_{i=1}^K \lambda_i \Psi_i(x) \quad \forall x \in X \}, \quad (4.22)$$

where sup is sought over $\lambda \stackrel{\text{def}}{=} \{\lambda_i\} \in \mathbb{R}^K$. A solution of the problem (4.22), that is $\lambda^ \stackrel{\text{def}}{=} \{\lambda_i^*\} \in \mathbb{R}^K$ such that*

$$\sigma^* \leq \Psi_0(x) + \sum_{i=1}^K \lambda_i^* \Psi_i(x) \quad \forall x \in X \quad (4.23)$$

exists if the inequality

$$0 \leq \sum_{i=1}^K v_i \Psi_i(x) \quad \forall x \in X \quad (4.24)$$

is valid only with $v_i = 0$, $i = 1, \dots, K$.

Proof. See the proof of Lemma 3.10 in [48]. \square

Proof of Proposition 4.4. By Lemma 4.5, to prove the Proposition 4.4(i), it is sufficient to show that, under Assumption 4.2, the inequality

$$0 \leq \sum_{i=1}^N v_i [\nabla \psi_i(z)^T \tilde{g}(\mu, z)] \quad \forall (\mu, z) \in F_M \quad (4.25)$$

can be valid only with $v_i = 0$, $i = 1, \dots, N$. Let us assume that (4.25) is valid and let us rewrite it in the form

$$0 \leq \nabla \psi(z)^T \tilde{g}(\mu, z) \quad \forall (\mu, z) \in F_M, \quad \text{where} \quad \psi(z) \stackrel{\text{def}}{=} \sum_{i=1}^N v_i \psi_i(z). \quad (4.26)$$

Let $z', z'' \in Z^0$ and let an admissible pair $(\mu(\cdot), z(\cdot))$ of the averaged system be such that $z(0) = z'$ and $z(\mathcal{T}) = z''$ for some $\mathcal{T} > 0$. Since $(\mu(t), z(t)) \in F \quad \forall t \in [0, \mathcal{T}]$ and since $F \subset F_M$ (see (4.5)), from (4.26) it follows that

$$\psi(z'') - \psi(z') = \int_0^{\mathcal{T}} \nabla \psi(z(t))^T \tilde{g}(\mu(t), z(t)) dt \geq 0 \quad \Rightarrow \quad \psi(z'') \geq \psi(z').$$

Since z', z'' can be arbitrary points in Z^0 , it follows that

$$\psi(z) = \text{const} \quad \forall z \in Z^0 \quad \Rightarrow \quad \psi(z) = \text{const} \quad \forall z \in cl Z^0.$$

The latter implies that

$$\nabla \psi(z) = \sum_{i=1}^N v_i \nabla \psi_i(z) = 0 \quad \forall z \in int(cl Z^0),$$

which, in turn, implies that $v_i = 0$, $i = 1, \dots, N$ (due to linear independence of $\nabla \psi_i(\cdot)$). The statement (i) of the proposition is proved. The proof of the statement (ii) is similar (see also the proof of Proposition 3.9 in [48]). \square

5. Construction of near optimal ACG families. Let us assume that, for any N and M , a solution $\zeta^{N,M}(z)$ of the (N, M) -approximating averaged dual problem exists and a solution $\eta_z^{N,M}(y)$ of the (N, M) -approximating associated problem exists for any $z \in Z$ (as follows from Proposition 4.4, these exist if Assumptions 4.2 and 4.3 are satisfied).

Define a control $u^{N,M}(y, z)$ as an optimal solution of the problem

$$\min_{u \in U} \{G(u, y, z) + \nabla \zeta^{N,M}(z)^T g(u, y, z) + \nabla \eta_z^{N,M}(y)^T f(u, y, z)\}. \quad (5.1)$$

That is,

$$u^{N,M}(y, z) = \argmin_{u \in U} \{G(u, y, z) + \nabla \zeta^{N,M}(z)^T g(u, y, z) + \nabla \eta_z^{N,M}(y)^T f(u, y, z)\}. \quad (5.2)$$

Assume that the system

$$y'_z(\tau) = f(u^{N,M}(y_z(\tau), z), y_z(\tau), z), \quad y_z(0) = y \in Y, \quad (5.3)$$

has a unique solution $y_z^{N,M}(\tau) \in Y$. Below, we introduce assumptions under which it will be established that $(u_z^{N,M}(\cdot), y_z^{N,M}(\cdot))$, where $u_z^{N,M}(\tau) \stackrel{\text{def}}{=} u_z^{N,M}(y_z^{N,M}(\tau), z)$, is a near optimal ACG family (see Theorem 5.8).

ASSUMPTION 5.1. *The following conditions are satisfied:*

- (i) *The optimal solution p^* of the IDLP problem (2.29) is unique, and the equality (2.33) is valid.*
- (ii) *The optimal solution of the averaged problem (2.26) (that is, an admissible pair $(\mu^*(\cdot), z^*(\cdot))$ that delivers minimum in (2.26)) exists and, for any continuous function $h(\mu, z) : F \rightarrow \mathbb{R}^1$, there exists a limit*

$$\lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{h}(\mu^*(t), z^*(t)) dt. \quad (5.4)$$

(iii) For almost all $t \in [0, \infty)$ and any $r > 0$, the p^* -measure of the set

$$\mathcal{B}_r(\mu^*(t), z^*(t)) \stackrel{\text{def}}{=} \{(\mu, z) : \rho(\mu, \mu^*(t)) + \|z - z^*(t)\| < r\}$$

is not zero. That is,

$$p^*(\mathcal{B}_r(\mu^*(t), z^*(t))) > 0. \quad (5.5)$$

Note that from Assumption 5.1(ii) it follows that the pair $(\mu^*(\cdot), z^*(\cdot))$ generates an occupational measure and from Assumptions 5.1(i) it follows that this measure coincides with p^* (see Corollary 3.4). That is,

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \tilde{h}(\mu^*(t), z^*(t)) dt = \int_F \tilde{h}(\mu, z) p^*(d\mu, dz). \quad (5.6)$$

The following statement gives sufficient conditions for the validity of Assumption 5.1(iii).

PROPOSITION 5.2. *Let Assumptions 5.1(i) and 5.1(ii) be satisfied. Then Assumption 5.1(iii) will be satisfied if the pair $(\mu^*(\cdot), z^*(\cdot))$ is \tilde{T} -periodic (\tilde{T} is some positive number) and if $\mu^*(\cdot)$ is piecewise continuous on $[0, \tilde{T}]$.*

Proof. Let t be a continuity point of $\mu^*(\cdot)$. Due to the assumed periodicity of the pair $(\mu^*(\cdot), z^*(\cdot))$,

$$\frac{1}{\tilde{T}} \int_0^{\tilde{T}} \tilde{h}(\mu^*(t), z^*(t)) dt = \int_F \tilde{h}(\mu, z) p^*(d\mu, dz)$$

and

$$p^*(\mathcal{B}_r(\mu^*(t), z^*(t))) = \frac{1}{\tilde{T}} \text{meas}\{t' : t' \in [0, \tilde{T}], (\mu^*(t'), z^*(t')) \in \mathcal{B}_r(\mu^*(t), z^*(t))\}. \quad (5.7)$$

Since t is a continuity point of $\mu^*(\cdot)$ and since $z^*(\cdot)$ is continuous, there exists $\alpha > 0$ such that $(\mu^*(t'), z^*(t')) \in \mathcal{B}_r(\mu^*(t), z^*(t)) \quad \forall t' \in [t - \alpha, t + \alpha]$. Hence, the right-hand-side in (5.7) is greater than $\frac{2\alpha}{\tilde{T}}$. This proves the required statement as the number of discontinuity points of $\mu^*(\cdot)$ is finite (due to the assumed piecewise continuity). \square

ASSUMPTION 5.3.

(i) For almost all $t \in [0, \infty)$, there exists an admissible pair $(u_t^*(\tau), y_t^*(\tau))$ of the associated system (considered with $z = z^*(t)$) such that $\mu^*(t)$ is the occupational measure generated by this pair on the interval $[0, \infty)$. That is, for any continuous $h(u, y)$,

$$\lim_{S \rightarrow \infty} S^{-1} \int_0^S h(u_t^*(\tau), y_t^*(\tau)) d\tau = \int_{U \times Y} h(u, y) \mu^*(t)(du, dy). \quad (5.8)$$

(ii) For almost all $t \in [0, \infty)$, for almost all $\tau \in [0, \infty)$ and for any $r > 0$, the $\mu^*(t)$ -measure of the set

$$\mathcal{B}_r(u_t^*(\tau), y_t^*(\tau)) \stackrel{\text{def}}{=} \{(u, y) : \|u - u_t^*(\tau)\| + \|y - y_t^*(\tau)\| < r\}$$

is not zero. That is,

$$\mu^*(t)(\mathcal{B}_r(u_t^*(\tau), y_t^*(\tau))) > 0. \quad (5.9)$$

The following proposition gives sufficient conditions for the validity of Assumption 5.3(ii).

PROPOSITION 5.4. *Let Assumptions 5.3(i) be valid. Then Assumption 5.3(ii) will be satisfied if, for almost all $t \in [0, \infty)$, the pair $(u_t^*(\tau), y_t^*(\tau))$ is T_t -periodic (T_t is some positive number) and if $u^*(\cdot)$ is piecewise continuous on $[0, T_t]$.*

Proof. The proof is similar to that of Proposition 5.2 \square

ASSUMPTION 5.5.

(i) The pair $(u_z^{N,M}(\tau), y_z^{N,M}(\tau))$, where $y_z^{N,M}(\tau)$ is the solution of (5.3) and $u_z^{N,M}(\tau) = u_z^{N,M}(y_z^{N,M}(\tau), z)$ is an ACG family that generates the occupational measure $\mu^{N,M}(du, dy|z)$ on the interval $[0, \infty)$, the latter

being independent of the initial conditions $y_z^{N,M}(0) = y$ for y in a neighbourhood of $y_t^*(\cdot)$. Also, for any continuous $h(u, y, z) : U \times Y \times Z \rightarrow \mathbb{R}^1$,

$$|S^{-1} \int_0^S h(u_z^{N,M}(\tau), y_z^{N,M}(\tau), z) d\tau - \int_{U \times Y} h(u, y, z) \mu^{N,M}(du, dy|z)| \leq \phi_h(S) \quad \forall z \in Z, \quad \lim_{S \rightarrow \infty} \phi_h(S) = 0. \quad (5.10)$$

(ii) The admissible pair of the averaged system $(\mu^{N,M}(\cdot), z^{N,M}(\cdot))$ generated by $(u_z^{N,M}(\cdot), y_z^{N,M}(\cdot))$ generates the occupational measure $\lambda^{N,M} \in \mathcal{P}(F)$, the latter being independent of the initial conditions $z^{N,M}(0) = z$ for z in a neighbourhood of $z^*(\cdot)$. Also, for any continuous function $\tilde{h}(\mu, z) : F \rightarrow \mathbb{R}^1$,

$$|\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \tilde{h}(\mu^{N,M}(t), z^{N,M}(t)) dt - \int_F \tilde{h}(\mu, z) \lambda^{N,M}(d\mu, dz)| \leq \phi_{\tilde{h}}(T), \quad \lim_{T \rightarrow \infty} \phi_{\tilde{h}}(T) = 0. \quad (5.11)$$

To state our next assumption, let us re-denote the occupational measure $\mu^{N,M}(du, dy|z)$ (introduced in Assumption 5.5 above) as $\mu^{N,M}(z)$ (that is, $\mu^{N,M}(du, dy|z) = \mu^{N,M}(z)$).

ASSUMPTION 5.6. For almost all $t \in [0, \infty)$, there exists an open ball $Q_t \subset \mathbb{R}^n$ centered at $z^*(t)$ such that:

(i) The occupational measure $\mu^{N,M}(z)$ is continuous on Q_t . Namely, for any $z', z'' \in Q_t$,

$$\rho(\mu^{N,M}(z'), \mu^{N,M}(z'')) \leq \kappa(\|z' - z''\|), \quad (5.12)$$

where $\kappa(\theta)$ is a function tending to zero when θ tends to zero ($\lim_{\theta \rightarrow 0} \kappa(\theta) = 0$). Also, for any $z', z'' \in Q_t$,

$$\left\| \int_{U \times Y} g(u, y, z') \mu^{N,M}(z')(du, dy) - \int_{U \times Y} g(u, y, z'') \mu^{N,M}(z'')(du, dy) \right\| \leq L \|z' - z''\|, \quad (5.13)$$

where L is a constant.

(ii) Let $z^{N,M}(\cdot)$ be the solution of the system

$$z'(t) = \tilde{g}(\mu^{N,M}(z(t)), z(t)), \quad z(0) = z^*(0). \quad (5.14)$$

We assume that, for any $t > 0$,

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \text{meas}\{A_t(N, M)\} = 0, \quad (5.15)$$

where

$$A_t(N, M) \stackrel{\text{def}}{=} \{t' \in [0, t] : z^{N,M}(t') \notin Q_{t'}\}. \quad (5.16)$$

In addition to the assumptions above, let us also introduce

ASSUMPTION 5.7. For each $t \in [0, \infty)$ such that $Q_t \neq \emptyset$, the following conditions are satisfied:

(i) For almost all $\tau \in [0, \infty)$, there exists an open ball $B_{t,\tau} \subset \mathbb{R}^m$ centered at $y_t^*(\tau)$ such that $u^{N,M}(y, z)$ is uniquely defined (the problem (5.1) has a unique solution) for $(y, z) \in B_{t,\tau} \times Q_t$.

(ii) The function $u^{N,M}(y, z)$ satisfies Lipschitz conditions on $B_{t,\tau} \times Q_t$. That is,

$$\|u^{N,M}(y', z') - u^{N,M}(y'', z'')\| \leq L(\|y' - y''\| + \|z' - z''\|) \quad \forall (y', z'), (y'', z'') \in B_{t,\tau} \times Q_t, \quad (5.17)$$

where L is a constant.

(iii) Let $y_t^{N,M}(\tau) \stackrel{\text{def}}{=} y_{z^*(t)}^{N,M}(\tau)$ be the solution of the system (5.3) considered with $z = z^*(t)$ and with the initial condition $y_z(0) = y_t^*(0)$. We assume that, for any $\tau > 0$,

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \text{meas}\{P_{t,\tau}(N, M)\} = 0, \quad (5.18)$$

where

$$P_{t,\tau}(N, M) \stackrel{\text{def}}{=} \{\tau' \in [0, \tau] : y_t^{N,M}(\tau') \notin B_{t,\tau'}\}. \quad (5.19)$$

THEOREM 5.8. *Let Assumptions 5.1, 5.3, 5.5, 5.6 and 5.7 be satisfied. Then the family $(u_z^{N,M}(\cdot), y_z^{N,M}(\cdot))$ introduced in Assumption 5.5(i) is a $\beta(N, M)$ -near optimal ACG family, with*

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \beta(N, M) = 0. \quad (5.20)$$

Also,

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \rho(\lambda^{N,M}, p^*) = 0, \quad (5.21)$$

where $\lambda^{N,M}$ is defined by (5.11).

Proof. The proof is given in Section 7. It is based on Lemma 5.9 stated at the end of this section. Note that in the process of the proof of the theorem it is established that

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \max_{t' \in [0, t]} \|z^{N,M}(t') - z^*(t')\| = 0 \quad \forall t \in [0, \infty), \quad (5.22)$$

where $z^{N,M}(\cdot)$ is the solution of (5.14). Also, it is shown that

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \rho(\mu^{N,M}(z^{N,M}(t)), \mu^*(t)) = 0 \quad (5.23)$$

for almost all $t \in [0, \infty)$, and

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} |\tilde{V}^{N,M} - \tilde{G}^*| = 0, \quad (5.24)$$

where

$$\tilde{V}^{N,M} \stackrel{\text{def}}{=} \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \tilde{G}(\mu^{N,M}(z^{N,M}(t)), z^{N,M}(t)) dt. \quad (5.25)$$

The relationship (5.24) implies the statement of the theorem with

$$\beta(N, M) \stackrel{\text{def}}{=} \tilde{V}^{N,M} - \tilde{G}^* \quad (5.26)$$

(see Definition 3.3). \square

LEMMA 5.9. *Let the assumptions of Theorem 5.8 be satisfied and let $t \in [0, \infty)$ be such that $Q_t \neq \emptyset$. Then*

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \max_{\tau' \in [0, \tau]} \|y_t^{N,M}(\tau') - y_t^*(\tau')\| = 0 \quad \forall \tau \in [0, \infty). \quad (5.27)$$

Also,

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \|u^{N,M}(y_t^{N,M}(\tau), z^*(t)) - u_t^*(\tau)\| = 0 \quad (5.28)$$

for almost all $\tau \in [0, \infty)$.

Proof. The proof is given in Section 7. \square

REMARK 5.10. *Note that from (4.12) and (5.21) it follows that*

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \rho(\lambda^{N,M}, p^{N,M}) = 0, \quad (5.29)$$

where $p^{N,M}$ is an arbitrary optimal solution of the (N, M) -approximating averaged problem (4.4).

6. Construction of asymptotically near optimal controls of the SP problem. In the previous section we discussed a construction of near optimal ACG families. In this section we will discuss a way how the latter can be used for a construction of asymptotically near optimal controls of the SP problem.

DEFINITION 6.1. *A control $u_\epsilon(\cdot)$ will be called asymptotically α -near optimal ($\alpha > 0$) in the SP problem (2.5) if the solution $(y_\epsilon(\cdot), z_\epsilon(\cdot))$ of the system (2.1)-(2.2) obtained with this control satisfies (2.4) (that is the triplet $(u_\epsilon(\cdot), y_\epsilon(\cdot), z_\epsilon(\cdot))$ is admissible) and if*

$$\lim_{\epsilon \rightarrow 0} \liminf_{T \rightarrow \infty} \frac{1}{T} \int_0^T G(u_\epsilon(t), y_\epsilon(t), z_\epsilon(t)) dt \leq \liminf_{\epsilon \rightarrow 0} V^*(\epsilon) + \alpha. \quad (6.1)$$

For simplicity, we will be dealing with a special case when $f(u, y, z) = f(u, y)$. That is, the right hand side in (2.1) is independent of z (the SP systems that have such a property are called “weakly coupled”). Note that in this case the set $W(z)$ defined in (2.15) does not depend on z too. That is, $W(z) = W$.

Let us also introduce the following assumptions about the functions $f(u, y)$ and $\tilde{g}(\mu, z)$.

ASSUMPTION 6.2.

(i) *There exists a positive definite matrix A_1 such that its eigenvalues are greater than a positive constant on Z and such that*

$$\begin{aligned} & (f(u, y') - f(u, y''))^T A_1 (y' - y'') \\ & \leq -(y' - y'')^T (y' - y'') \quad \forall y', y'' \in \mathbb{R}^m, \quad \forall u \in U. \end{aligned} \quad (6.2)$$

(ii) *There exists a positive definite matrix A_2 such that its eigenvalues are greater than a positive constant on Z and such that*

$$\begin{aligned} & (\tilde{g}(\mu, z') - \tilde{g}(\mu, z''))^T A_2 (z' - z'') \\ & \leq -(z' - z'')^T (z' - z'') \quad \forall z', z'' \in \mathbb{R}^n, \quad \forall \mu \in W. \end{aligned} \quad (6.3)$$

Note that these are Liapunov type stability conditions and, as has been established in [44], their fulfillment is sufficient for the validity of the statement that the SP system is uniformly approximated by the averaged system (see Definition 2.4). Also, as can be readily verified, Assumption 6.2(i) implies that the solutions $y(\tau, u(\cdot), y_1)$ and $y(\tau, u(\cdot), y_2)$ of the associated system (2.11) obtained with an arbitrary control $u(\cdot)$ and with initial values $y(0) = y_1$ and $y(0) = y_2$ (y_1 and y_2 being arbitrary vectors in Y) satisfy the inequality

$$\|y(\tau, u(\cdot), y_1) - y(\tau, u(\cdot), y_2)\| \leq c_1 e^{-c_2 \tau} \|y_1 - y_2\|, \quad (6.4)$$

where c_1, c_2 are some positive constants. Similarly, Assumption 6.2(ii) implies that the solutions $z(t, \mu(\cdot), z_1)$ and $z(t, \mu(\cdot), z_2)$ of the averaged system (2.18) obtained with an arbitrary control $\mu(\cdot)$ and with initial values $z(0) = z_1$ and $z(0) = z_2$ (z_1 and z_2 being arbitrary vectors in Z) satisfy the inequality

$$\|z(t, \mu(\cdot), z_1) - z(t, \mu(\cdot), z_2)\| \leq c_3 e^{-c_4 t} \|z_1 - z_2\|, \quad (6.5)$$

where c_3, c_4 are some positive constants.

From the validity of (6.4) and (6.5) it follows that the associated system (2.11) and the averaged system (2.18) have unique forward invariant sets which also are global attractors for the solutions of these systems (see Theorem 3.1(ii) in [39]). For simplicity, we will assume that Y and Z are these sets.

Let $(u_z^{N,M}(\tau), y_z^{N,M}(\tau))$ be the ACG family introduced in Assumptions 5.5(i) and let $\mu^{N,M}(du, dy|z) = \mu^{N,M}(z)$, $z^{N,M}(t)$ and $\mu^{N,M}(z^{N,M}(t))$ be generated by this family as assumed in Section 5 (all the assumptions made in that section are supposed to be satisfied in the consideration below). Let $y_z^{N,M}(\tau, y)$ stand for the solution of the associated system (2.11) obtained with the control $u_z^{N,M}(\tau)$ and with the initial condition $y_z^{N,M}(0, y) = y \in Y$. From (6.4) it follows that

$$\|y_z^{N,M}(\tau, y) - y_z^{N,M}(\tau, y')\| \leq c_1 e^{-c_2 \tau} \max_{y', y'' \in Y} \|y' - y''\|.$$

The latter implies that, for any Lipschitz continuous function $h(u, y, z)$, there exists $\bar{\phi}_h(S)$, $\lim_{S \rightarrow \infty} \bar{\phi}_h(S) = 0$, such that

$$|S^{-1} \int_0^S h(u_z^{N,M}(\tau), y_z^{N,M}(\tau, y), z) d\tau - S^{-1} \int_0^S h(u_z^{N,M}(\tau), y_z^{N,M}(\tau), z) d\tau| \leq \bar{\phi}_h(S), \quad (6.6)$$

which, due to (5.10), implies that

$$|S^{-1} \int_0^S h(u_z^{N,M}(\tau), y_z^{N,M}(\tau, y), z) d\tau - \int_{U \times Y} h(u, y, z) \mu^{N,M}(du, dy|z)| \leq \phi_h(S) + \bar{\phi}_h(S) \stackrel{\text{def}}{=} \bar{\bar{\phi}}_h(S), \quad (6.7)$$

with $\lim_{S \rightarrow \infty} \bar{\bar{\phi}}_h(S) = 0$. Hence,

$$\lim_{S \rightarrow \infty} \rho(\mu^{N,M}(S, y), \mu^{N,M}(z)) \leq \phi(S), \quad \lim_{S \rightarrow \infty} \phi(S) = 0, \quad (6.8)$$

where $\mu^{N,M}(S, y)$ is the occupational measure generated by the pair $(u_z^{N,M}(\tau), y_z^{N,M}(\tau, y))$ on the interval $[0, S]$. That is, the family of measures $\mu^{N,M}(z)$ is uniformly attainable by the associated system with the use of the control $u_z^{N,M}(\tau)$ (see Definition 4.3 in [48]).

Partition the interval $[0, \infty)$ by the points

$$t_l = l\Delta(\epsilon), \quad l = 0, 1, \dots, \quad (6.9)$$

where $\Delta(\epsilon) > 0$ is such that

$$\lim_{\epsilon \rightarrow 0} \Delta(\epsilon) = 0, \quad \lim_{\epsilon \rightarrow 0} \frac{\Delta(\epsilon)}{\epsilon} = \infty. \quad (6.10)$$

Define the control $u_\epsilon^{N,M}(t)$ by the equation

$$u_\epsilon^{N,M}(t) \stackrel{\text{def}}{=} u_{z^{N,M}(t_l)}^{N,M}\left(\frac{t - t_l}{\epsilon}\right) \quad \forall t \in [t_l, t_{l+1}), \quad l = 0, 1, \dots \quad (6.11)$$

THEOREM 6.3. *Let the assumptions of Theorem 5.8 be satisfied and let the function*

$$\mu^{N,M}(t) \stackrel{\text{def}}{=} \mu^{N,M}(z^{N,M}(t)) \quad (6.12)$$

has the following piecewise continuity property: for any $\mathcal{T} > 0$, there may exist no more than a finite number of points $\mathcal{T}_i \in (0, \mathcal{T})$, $i = 1, \dots, k$, with

$$k \leq c\mathcal{T}, \quad c = \text{const}, \quad (6.13)$$

such that, for any $t \neq \mathcal{T}_i$,

$$\max\{|\tilde{g}(\mu^{N,M}(t'), z) - \tilde{g}(\mu^{N,M}(t), z)|, |\tilde{G}(\mu^{N,M}(t'), z) - \tilde{G}(\mu^{N,M}(t), z)|\} \leq \nu(t - t') \quad \forall t' \in (t - a_t, t + a_t) \quad (6.14)$$

where $\nu(\cdot)$ is monotone decreasing, with $\lim_{\theta \rightarrow 0} \nu(\theta) = 0$, and where $a_t > 0$, with r_δ ,

$$r_\delta \stackrel{\text{def}}{=} \inf\{a_t : t \notin \cup_{i=1}^k (\mathcal{T}_i - \delta, \mathcal{T}_i + \delta)\},$$

being a positive continuous function of δ (which may tend to zero when δ tends to zero). Let also Assumption 6.2 be valid and the solution $(y_\epsilon^{N,M}(\cdot), z_\epsilon^{N,M}(\cdot))$ of the system (2.1)-(2.2) obtained with the control $u_\epsilon^{N,M}(\cdot)$ and with the initial conditions $(y_\epsilon^{N,M}(0), z_\epsilon^{N,M}(0)) = (y^{N,M}(0), z^{N,M}(0))$ satisfies the inclusion (2.4). Then the control $u_\epsilon^{N,M}(\cdot)$ is $\beta(N, M)$ -asymptotically near optimal in the problem (2.5), where $\beta(N, M)$ is defined in (5.26). Also,

$$\lim_{\epsilon \rightarrow 0} \sup_{t \in [0, \infty)} \|z_\epsilon^{N,M}(t) - z^{N,M}(t)\| = 0 \quad (6.15)$$

and, if the triplet $(u_\epsilon^{N,M}(\cdot), y_\epsilon^{N,M}(\cdot), z_\epsilon^{N,M}(\cdot))$ generates the occupational measure $\gamma_\epsilon^{N,M}$ on the interval $[0, \infty)$ (see (1.3)), then

$$\rho(\gamma_\epsilon^{N,M}, \Phi(\lambda^{N,M})) \leq \kappa(\epsilon) \quad \text{where} \quad \lim_{\epsilon \rightarrow 0} \kappa(\epsilon) = 0, \quad (6.16)$$

with $\lambda^{N,M}$ being the occupational measure generated by $(\mu^{N,M}(\cdot), z^{N,M}(\cdot))$ (see (5.11)) and the map $\Phi(\cdot)$ being defined by (2.40).

Proof. The proof of the theorem is given in Section 8. \square

Note that from (5.29) and (6.16) it follows (due to continuity of $\Phi(\cdot)$) that

$$\rho(\gamma_\epsilon^{N,M}, \Phi(p^{N,M})) \leq \kappa(\epsilon) + \theta(N, M), \quad \text{where} \quad \lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \theta(N, M) = 0 \quad (6.17)$$

and where $p^{N,M}$ is an arbitrary optimal solution of the (N, M) -approximating averaged problem (4.4). This problem always has an optimal solution that can be presented in the form (see Section 7 below)

$$p^{N,M} \stackrel{\text{def}}{=} \sum_{k=1}^K p_k \delta_{(\mu_k, z_k)}, \quad \sum_{k=1}^K p_k = 1, \quad p_k > 0, \quad k = 1, \dots, K, \quad (6.18)$$

where

$$\mu_k = \sum_{j=1}^{J_k} q_j^k \delta_{(u_j^k, y_j^k)}, \quad \sum_{j=1}^{J_k} q_j^k = 1, \quad q_j^k > 0, \quad j = 1, \dots, J_k, \quad (6.19)$$

$\delta_{(u_j^k, y_j^k)}$ being the Dirac measure concentrated at $(u_j^k, y_j^k) \in U \times Y$ ($j = 1, \dots, J_k$) and $\delta_{(\mu_k, z_k)}$ being the Dirac measure concentrated at $(\mu_k, z_k) \in \mathcal{P}(U \times Y) \times Z$ ($k = 1, \dots, K$). As can be readily verified,

$$\Phi\left(\sum_{k=1}^K p_k \delta_{(\mu_k, z_k)}\right) = \sum_{k=1}^K \sum_{j=1}^{J_k} p_k q_j^k \delta_{(u_j^k, y_j^k, z_k)}, \quad (6.20)$$

where $\delta_{(u_j^k, y_j^k, z_k)}$ is the Dirac measure concentrated at (u_j^k, y_j^k, z_k) . Thus, by (6.17),

$$\rho(\gamma_\epsilon^{N,M}, \sum_{k=1}^K \sum_{j=1}^{J_k} p_k q_j^k \delta_{(u_j^k, y_j^k, z_k)}) \leq \kappa(\epsilon) + \theta(N, M). \quad (6.21)$$

That is, for N, M large enough and ϵ small enough, the occupational measure $\gamma_\epsilon^{N,M}$ is approximated by a convex combination of the Dirac measures, which implies, in particular, that the state trajectory $(y_\epsilon^{N,M}(\cdot), z_\epsilon^{N,M}(\cdot))$ spends a non-zero proportion of time in a vicinity of each of the points (u_j^k, y_j^k, z_k) .

To illustrate the construction of asymptotically near optimal controls, let us conclude this section with an example (which was briefly discussed in [48]). Consider the optimal control problem

$$\inf_{(u(\cdot), y_\epsilon(\cdot), z_\epsilon(\cdot))} \liminf_{T \rightarrow \infty} \frac{1}{T} \int_0^T (0.1u_1^2(t) + 0.1u_2^2(t) - z_1^2(t)) dt = V^*(\epsilon), \quad (6.22)$$

where minimization is over the controls $u(\cdot) = (u_1(\cdot), u_2(\cdot))$,

$$(u_1(t), u_2(t)) \in U \stackrel{\text{def}}{=} \{(u_1, u_2) : |u_i| \leq 1, \quad i = 1, 2\}, \quad (6.23)$$

and the corresponding solutions $y_\epsilon(\cdot) = (y_{1,\epsilon}(\cdot), y_{2,\epsilon}(\cdot))$ and $z_\epsilon(\cdot) = (z_{1,\epsilon}(\cdot), z_{2,\epsilon}(\cdot))$ of the SP system

$$\epsilon y_i'(t) = -y_i(t) + u_i(t), \quad i = 1, 2, \quad (6.24)$$

$$z_1'(t) = z_2(t), \quad z_2'(t) = -4z_1(t) - 0.3z_2(t) - y_1(t)u_2(t) + y_2(t)u_1(t), \quad (6.25)$$

with

$$(y_1(t), y_2(t)) \in Y \stackrel{\text{def}}{=} \{(y_1, y_2) : |y_i| \leq 1, \quad i = 1, 2\}$$

and with

$$(z_1(t), z_2(t)) \in Z \stackrel{\text{def}}{=} \{(z_1, z_2) : |z_1| \leq 2.5, \quad |z_2| \leq 4.5\}.$$

The averaged system (2.18) takes in this case the form

$$z_1'(t) = z_2(t), \quad z_2'(t) = -4z_1(t) + \int_{U \times Y} (-y_1 u_2 + y_2 u_1) \mu(t)(du, dy) \quad (6.26)$$

where

$$\mu(t) \in W \stackrel{\text{def}}{=} \{\mu \in \mathcal{P}(U \times Y) : \int_{U \times Y} [\frac{\partial \phi(y)}{\partial y_1} (-y_1 + u_1) + \frac{\partial \phi(y)}{\partial y_2} (-y_2 + u_2)] \mu(du, dy) = 0 \quad \forall \phi(\cdot) \in C^1(\mathbb{R}^2)\}, \quad (6.27)$$

Note that, as can be readily verified, the function $f(u, y) = (-y_1 + u_1, -y_2 + u_2)$ satisfies Assumption 6.2(i) and the function $\tilde{g}(\mu, z) = (z_2, -4z_1 + \int_{U \times Y} (-y_1 u_2 + y_2 u_1) \mu(du, dy))$ satisfies Assumption 6.2(ii).

The (N, M) -approximating averaged problem (4.4) was constructed in this example with the use of the monomials $z_1^{j_1} z_2^{j_2}$ ($1 \leq j_1 + j_2 \leq 5$) as the test functions in (4.3) and the monomials $y_1^{i_1} y_2^{i_2}$ ($1 \leq i_1 + i_2 \leq 5$) as the test functions in (4.1). Note that $N, M = 35$ in this case (recall that N stands for the number of constraints in (4.3) and M stands for the number of constraints in (4.1)). This problem was solved numerically with the help of the linear programming based algorithm described in Section 4.3 of [48], its output including the optimal value of the problem, an optimal solution of the problem and solutions of the corresponding averaged and associated dual problems.

The optimal value of the problem was obtained to be approximately equal to -1.186 :

$$\tilde{G}^{35,35} \approx -1.186. \quad (6.28)$$

Along with the optimal value, the points

$$z_k = (z_{1,k}, z_{2,k}) \in Z, \quad k = 1, \dots, K, \quad (6.29)$$

and weights $\{p_k\}$ that enter the expansion (6.18) as well as the points

$$u_j^k = (u_{1,j}^k, u_{2,j}^k) \in U, \quad y_j^k = (y_{1,j}^k, y_{2,j}^k) \in Y, \quad j = 1, \dots, J_k, \quad k = 1, \dots, K, \quad (6.30)$$

and the corresponding weights $\{q_j^k\}$ that enter the expansions (6.19) were numerically found. In Figure 1, the points $\{z_k\}$ that enter the expansion (6.18) are marked with dots on the “ z -plane”. Corresponding to each such a point z_k , there are points $\{y_j^k\}$ that enter the expansion (6.19). These points are marked with dots on the “ y -plane” in Figure 2 for $z_k \approx (1.07, -0.87)$ (which is one of the points marked in Fig.1; for other points marked in Fig. 1, the configurations of the corresponding $\{y_j^k\}$ points look similar).

The expansions (4.15) and (4.20) that define solutions of the (N, M) -approximating averaged and (N, M) -approximating associated dual problems take the form

$$\zeta^{35,35}(z) = \sum_{1 \leq j_1 + j_2 \leq 5} \lambda_{j_1, j_2}^{35,35} z_1^{j_1} z_2^{j_2}, \quad \eta_z^{35,35}(y) = \sum_{1 \leq i_1 + i_2 \leq 5} \omega_{z, i_1, i_2}^{35,35} y_1^{i_1} y_2^{i_2}, \quad (6.31)$$

where the coefficients $\{\lambda_{j_1, j_2}^{35,35}\}$ and $\{\omega_{z, i_1, i_2}^{35,35}\}$ are obtained as a part of the solution with the above mentioned algorithm. Using $\zeta^{35,35}(z)$ and $\eta_z^{35,35}(y)$, one can compose the problem (5.1):

$$\begin{aligned} \min_{u_i \in [-1, 1]} \{ & 0.1u_1^2 + 0.1u_2^2 - z_1^2 + \frac{\partial \zeta^{35,35}(z)}{\partial z_1} z_2 + \frac{\partial \zeta^{35,35}(z)}{\partial z_2} (-4z_1 - 0.3z_2 - y_1 u_2 + y_2 u_1) + \frac{\partial \eta_z^{35,35}(y)}{\partial y_1} (-y_1 + u_1) \\ & + \frac{\partial \eta_z^{35,35}(y)}{\partial y_2} (-y_2 + u_2) \}, \end{aligned} \quad (6.32)$$

the solution of which is written in the form

$$u_i^{35,35}(y, z) = \begin{cases} -5b_i^{35,35}(y, z) & if & |5b_i^{35,35}(y, z)| \leq 1, \\ -1 & if & -5b_i^{35,35}(y, z) < -1, \\ 1 & if & -5b_i^{35,35}(y, z) > 1, \end{cases} \quad i = 1, 2, \quad (6.33)$$

where $b_1^{35,35}(y, z) \stackrel{\text{def}}{=} \frac{\partial \zeta^{35,35}(z)}{\partial z_2} y_2 + \frac{\partial \eta_z^{35,35}(y)}{\partial y_1}$ and $b_2^{35,35}(y, z) \stackrel{\text{def}}{=} -\frac{\partial \zeta^{35,35}(z)}{\partial z_2} y_1 + \frac{\partial \eta_z^{35,35}(y)}{\partial y_2}$.

Using the feedback controls $u_i^{35,35}(y, z)$, $i = 1, 2$, with fixed $z = z_k \approx (1.07, -0.87)$ and integrating the associated system with MATLAB from the initial conditions defined by one of the points marked in Figure 2, one obtains a periodic solution $y_z^{35,35}(\tau) = (y_{1,z}^{35,35}(\tau), y_{2,z}^{35,35}(\tau))$. The corresponding square like state trajectory of the associated system is also depicted in Figure 2. Note that this trajectory is located in a close vicinity of the marked points, this being consistent with the comments made after the statement of Theorem 6.3.

Using the same controls $u_i^{35,35}(y, z)$, $i = 1, 2$, in the SP system (6.24)-(6.25) and integrating the latter (taken with $\epsilon = 0.01$ and $\epsilon = 0.001$) with MATLAB from the initial conditions defined by one of the points marked in Figure 1 and one of the points marked in Figure 2, one obtains visibly periodic solutions, the images of which

are depicted in Figures 3 and 4, with the state trajectory of the slow dynamics $z_\epsilon^{35,35}(t) = (z_{1\epsilon}^{35,35}(t), z_{2\epsilon}^{35,35}(t))$ being also depicted in Figure 1. The slow z -components appear to be moving periodically along an ellipse like figure on the plane (z_1, z_2) , with the period being approximately equal to 3.16. Note that this figure and the period appear to be the same for $\epsilon = 0.01$ and for $\epsilon = 0.001$, with the marked points being located on or very close to the ellipse like figure in Fig. 1 (which again is consistent with the comments made after Theorem 6.3). In Figures 3 and 4, the fast y -components are moving along square like figures (similar to that in Fig. 1) centered around the points on the “ellipse”, with about 50 rounds for the case $\epsilon = 0.01$ (Fig. 3) and about 500 rounds for the case $\epsilon = 0.001$ (Fig. 4). The values of the objective functions obtained for these two cases are approximately the same and ≈ -1.177 , the latter being close to the value of $\tilde{G}^{35,35}$ (see (6.28)). Due to (2.34) and due to (4.8), this indicates that the found solution is close to the optimal one.

Graphs of $z_\epsilon^{35,35}(t)$ and $y_z^{35,35}(\tau)$

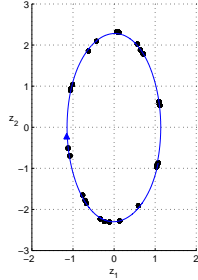


Fig. 1: $z_\epsilon^{35,35}(t) = (z_{1\epsilon}^{35,35}(t), z_{2\epsilon}^{35,35}(t))$

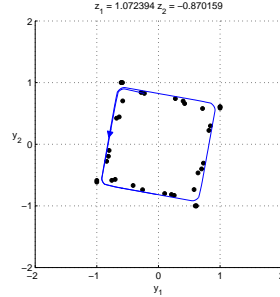


Fig.2: $y_z^{35,35}(\tau) = (y_{1,z}^{35,35}(\tau), y_{2,z}^{35,35}(\tau))$

Images of the state trajectories of the SP system for $\epsilon = 0.01$ and $\epsilon = 0.001$

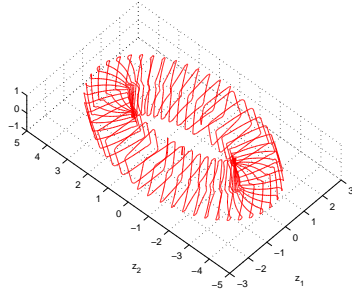


Fig. 3: $(y_\epsilon^{35,35}(t), z_\epsilon^{35,35}(t))$ for $\epsilon = 0.01$

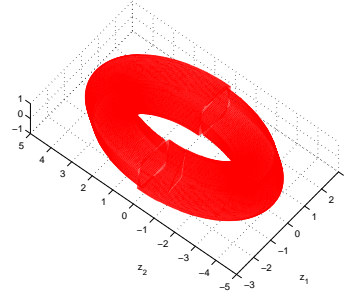


Fig.4: $(y_\epsilon^{35,35}(t), z_\epsilon^{35,35}(t))$ for $\epsilon = 0.001$

Note, in conclusion, that by taking $\epsilon = 0$ in (6.24), one obtains $y_i(t) = u_i(t)$, $i = 1, 2$, and, thus, arrives at the equality

$$-y_1(t)u_2(t) + y_2(t)u_1(t) = 0 \quad \forall t,$$

which makes the slow dynamics uncontrolled and leads to the optimality of the “trivial” steady state regime: $u_1(t) = u_2(t) = y_1(t) = y_2(t) = z_1(t) = z_2(t) = 0 \quad \forall t$ implying that $V^*(0) = 0$. Thus, in the present example

$$\lim_{\epsilon \rightarrow 0} V^*(\epsilon) \approx -1.177 < 0 = V^*(0).$$

7. Proof of Theorem 5.8. Note, first of all, that there exists an optimal solution $p^{N,M}$ of the problem (4.4) which is presented as a convex combination of (no more than $N+1$) Dirac measures (see, e.g., Theorems A.4 and A.5 in [69]). That is,

$$p^{N,M} = \sum_{k=1}^{K^{N,M}} p_k^{N,M} \delta_{(\mu_k^{N,M}, z_k^{N,M})}, \quad (7.1)$$

where $\delta_{(\mu_k^{N,M}, z_k^{N,M})}$ is the Dirac measure concentrated at $(\mu_k^{N,M}, z_k^{N,M})$ and

$$(\mu_k^{N,M}, z_k^{N,M}) \in F_M, \quad p_k^{N,M} > 0, \quad k = 1, \dots, K^{N,M} \leq N+1; \quad \sum_{k=1}^{K^{N,M}} p_k^{N,M} = 1. \quad (7.2)$$

LEMMA 7.1. For any $k = 1, \dots, K^{N,M}$,

$$\mu_k^{N,M} = \operatorname{argmin}_{\mu \in W_M(z_k^{N,M})} \{ \tilde{G}(\mu, z_k^{N,M}) + \nabla \zeta^{N,M}(z_k^{N,M})^T \tilde{g}(\mu, z_k^{N,M}) \}. \quad (7.3)$$

That is, $\mu_k^{N,M}$ is a minimizer of the problem

$$\min_{\mu \in W_M(z_k^{N,M})} \{ \tilde{G}(\mu, z_k^{N,M}) + \nabla \zeta^{N,M}(z_k^{N,M})^T \tilde{g}(\mu, z_k^{N,M}) \}. \quad (7.4)$$

Proof. The proof is similar to that of Lemma 5.1 in [48] \square

LEMMA 7.2. In the presentation (7.1) of an optimal solution $p^{N,M}$ of the problem (4.4), $\mu_k^{N,M}$ can be chosen as follows:

$$\mu_k^{N,M} = \sum_{j=1}^{J^{N,M,k}} q_j^{N,M,k} \delta_{(u_j^{N,M,k}, y_j^{N,M,k})}, \quad k = 1, \dots, K^{N,M}, \quad (7.5)$$

where

$$q_j^{N,M,k} > 0, \quad j = 1, \dots, J^{N,M,k}, \quad \sum_{j=1}^{J^{N,M,k}} q_j^{N,M,k} = 1, \quad (7.6)$$

and

$$J^{N,M,k} \leq N + M + 2. \quad (7.7)$$

In (7.5), $\delta_{(u_j^{N,M,k}, y_j^{N,M,k})} \in \mathcal{P}(U \times Y)$ are the Dirac measures concentrated at $(u_j^{N,M,k}, y_j^{N,M,k}) \in U \times Y$, $j = 1, \dots, J^{N,M,k}$, with

$$\begin{aligned} u_j^{N,M,k} = \operatorname{argmin}_{u \in U} \{ & G(u, y_j^{N,M,k}, z_k^{N,M}) + \nabla \zeta^{N,M}(z_k^{N,M})^T g(u, y_j^{N,M,k}, z_k^{N,M}) \\ & + \nabla \eta^{N,M}(y_j^{N,M,k})^T f(u, y_j^{N,M,k}, z_k^{N,M}) \}. \end{aligned} \quad (7.8)$$

Proof. The proof is similar to that of Lemma 5.2 in [48] \square

LEMMA 7.3. For any t such that (5.5) is satisfied, there exists a sequence

$$(\mu_{k_{N,M}}^{N,M}, z_{k_{N,M}}^{N,M}) \in \{(\mu_k^{N,M}, z_k^{N,M}), k = 1, \dots, K^{N,M}\}, \quad N = 1, 2, \dots, \quad M = 1, 2, \dots, \quad (7.9)$$

(with $\{(\mu_k^{N,M}, z_k^{N,M}), k = 1, \dots, K^{N,M}\}$ being the set of concentration points of the Dirac measures in (7.1)) such that

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} [\rho(\mu^*(t), \mu_{k_{N,M}}^{N,M}) + \|z^*(t) - z_{k_{N,M}}^{N,M}\|] = 0. \quad (7.10)$$

Let t be such that (7.10) is valid and let $(\mu_{k_{N,M}}^{N,M}, z_{k_{N,M}}^{N,M})$ be as in (7.10), then for any τ such that (5.9) is satisfied, there exists a sequence

$$(u_{j_{N,M}}^{N,M,k^{N,M}}, y_{j_{N,M}}^{N,M,k^{N,M}}) \in \{(u_j^{N,M,k^{N,M}}, y_j^{N,M,k^{N,M}}), j = 1, \dots, J^{N,M,k^{N,M}}\}, \quad N = 1, 2, \dots, \quad M = 1, 2, \dots, \quad (7.11)$$

($\{(u_j^{N,M,k^{N,M}}, y_j^{N,M,k^{N,M}}), j = 1, \dots, J^{N,M,k^{N,M}}\}$ being the set of concentration points of the Dirac measures in (7.5) taken with $k = k_{N,M}$) such that

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} [\|u_t^*(\tau) - u_{j_{N,M}}^{N,M,k^{N,M}}\| + \|y_t^*(\tau) - y_{j_{N,M}}^{N,M,k^{N,M}}\|] = 0. \quad (7.12)$$

Proof. Assume that (7.10) is not true. Then there exists $r > 0$ and sequences N_i , $M_{i,j}$ with $i = 1, \dots$, $j = 1, \dots$, and with $\lim_{i \rightarrow \infty} N_i = \infty$, $\lim_{j \rightarrow \infty} M_{i,j} = \infty$ such that

$$\text{dist}((\mu^*(t), z^*(t)), \Theta^{N_i, M_{i,j}}) \geq r, \quad (7.13)$$

where $\Theta^{N,M}$ is the set of the concentration points of the Dirac measures in (7.1), that is,

$$\Theta^{N,M} \stackrel{\text{def}}{=} \{(\mu_k^{N,M}, z_k^{N,M}), k = 1, \dots, K^{N,M}\},$$

taken with $N = N_i$ and $M = M_{i,j}$, and where

$$\text{dist}((\mu, z), \Theta^{N,M}) \stackrel{\text{def}}{=} \min_{(\mu', z') \in \Theta^{N,M}} \{\rho(\mu, \mu') + \|z - z'\|\}.$$

Hence,

$$(\mu_k^{N_i, M_{i,j}}, z_k^{N_i, M_{i,j}}) \notin \mathcal{B}_r(\mu^*(t), z^*(t)), \quad k = 1, \dots, K^{N_i, M_{i,j}}, \quad i, j = 1, 2, \dots.$$

The latter implies that

$$p^{N_i, M_{i,j}}(\mathcal{B}_r(\mu^*(t), z^*(t))) = 0, \quad i, j = 1, 2, \dots, \quad (7.14)$$

where $p^{N,N}$ is defined by (7.1). Due to the fact that the optimal solution p^* of the IDLP problem (2.29) is unique (Assumption 5.1(i)), the relationship (4.12) is valid. Consequently,

$$\lim_{i \rightarrow \infty} \limsup_{j \rightarrow \infty} \rho(p^{N_i, M_{i,j}}, p^*) = 0. \quad (7.15)$$

From (7.14) and (7.15) it follows that

$$p^*(\mathcal{B}_r(\mu^*(t), z^*(t))) \leq \lim_{i \rightarrow \infty} \limsup_{j \rightarrow \infty} p^{N_i, M_{i,j}}(\mathcal{B}_r(\mu^*(t), z^*(t))) = 0.$$

The latter contradicts to (5.5). Thus, (7.10) is proved.

Assume now that (7.12) is not valid. Then there exists $r > 0$ and sequences N_i , $M_{i,j}$ with $i = 1, \dots$, $j = 1, \dots$, and with $\lim_{i \rightarrow \infty} N_i = \infty$, $\lim_{j \rightarrow \infty} M_{i,j} = \infty$ such that

$$\text{dist}((u_t^*(\tau), y_t^*(\tau)), \theta^{N_i, M_{i,j}}) \geq r, \quad i, j = 1, 2, \dots, \quad (7.16)$$

where $\theta^{N,M}$ is the set of the concentration points of the Dirac measures in (7.5),

$$\theta^{N,M} \stackrel{\text{def}}{=} \{(u_j^{N,M, k^{N,M}}, y_j^{N,M, k^{N,M}}), \quad j = 1, \dots, J^{N,M, k^{N,M}}\},$$

taken with $k = k^{N,M}$ and with $N = N_i$, $M = M_{i,j}$, and where

$$\text{dist}((u, y), \theta^{N,M}) \stackrel{\text{def}}{=} \min_{(u', y') \in \theta^{N,M}} \{\|u - u'\| + \|y - y'\|\}.$$

From (7.16) it follows that

$$(u_j^{N_i, M_{i,j}, k^{N_i, M_{i,j}}}, y_j^{N_i, M_{i,j}, k^{N_i, M_{i,j}}}) \notin \mathcal{B}_r(u_t^*(\tau), y_t^*(\tau)), \quad j = 1, \dots, J^{N_i, M_{i,j}, k^{N_i, M_{i,j}}}, \quad i, j = 1, 2, \dots.$$

The latter implies that

$$\mu_{k^{N_i, M_{i,j}}}^{N_i, M_{i,j}}(\mathcal{B}_r(u_t^*(\tau), y_t^*(\tau))) = 0, \quad i, j = 1, 2, \dots. \quad (7.17)$$

where $\mu_k^{N,M}$ is defined by (7.5) (taken with $k = k^{N_i, M_{i,j}}$, $N = N_i$ and $M = M_{i,j}$).

From (7.10) it follows, in particular, that

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \rho(\mu^*(t), \mu_{k^{N,M}}^{N,M}) = 0 \quad \Rightarrow \quad \lim_{i \rightarrow \infty} \limsup_{j \rightarrow \infty} \rho(\mu^*(t), \mu_{k^{N_i, M_{i,j}}}^{N_i, M_{i,j}}) = 0. \quad (7.18)$$

The later and (7.17) lead to

$$\mu^*(t)(B_r(u_t^*(\tau), y_t^*(\tau))) \leq \lim_{i \rightarrow \infty} \limsup_{j \rightarrow \infty} \mu_{k^{N_i, M_i, j}}^{N_i, M_i, j}(B_r(u_t^*(\tau), y_t^*(\tau))) = 0,$$

which contradicts to (5.9). Thus (7.12) is proved. \square

LEMMA 7.4. *For any $t \in [0, \infty)$ such that Q_t is not empty and (5.5) is valid for an arbitrary $r > 0$, and for any $\tau \in [0, \infty)$ such that $B_{t, \tau}$ is not empty and (5.9) is valid for an arbitrary $r > 0$,*

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \|u_t^*(\tau) - u^{N, M}(y_t^*(\tau), z^*(t))\| = 0. \quad (7.19)$$

Proof. By Lemma 7.3, there exist $(\mu_{k^{N, M}}^{N, M}, z_{k^{N, M}}^{N, M})$ such that (7.10) is satisfied and there exist $(u_{j^{N, M}}^{N, M, k^{N, M}}, y_{j^{N, M}}^{N, M, k^{N, M}})$ such that (7.12) is satisfied.

Note that, due to (7.8),

$$u_{j^{N, M}}^{N, M, k^{N, M}} = u(y_{j^{N, M}}^{N, M, k^{N, M}}, z_{k^{N, M}}^{N, M}), \quad (7.20)$$

where $u(y, z)$ is as in (5.2). From (7.10) and (7.12) it follows that $z_{k^{N, M}}^{N, M} \in Q_t$ and $y_{j^{N, M}}^{N, M, k^{N, M}} \in B_{t, \tau}$ for N and M large enough. Hence, one can use (5.17) to obtain

$$\begin{aligned} \|u_t^*(\tau) - u^{N, M}(y_t^*(\tau), z^*(t))\| &\leq \|u_t^*(\tau) - u_{j^{N, M}}^{N, M, k^{N, M}}\| + \|u(y_{j^{N, M}}^{N, M, k^{N, M}}, z_{k^{N, M}}^{N, M}) - u^{N, M}(y_t^*(\tau), z^*(t))\| \\ &\leq \|u_t^*(\tau) - u_{j^{N, M}}^{N, M, k^{N, M}}\| + L(\|y_t^*(\tau) - y_{j^{N, M}}^{N, M, k^{N, M}}\| + \|z^*(t) - z_{k^{N, M}}^{N, M}\|). \end{aligned}$$

By (7.10) and (7.12), the latter implies (7.19). \square

Proof of Lemma 5.9. Let $t \in [0, \infty)$ be such that Q_t is not empty and (5.5) is satisfied for an arbitrary $r > 0$. Note that, from the assumptions made, it follows that (7.19) is valid for almost all $\tau \in [0, \infty)$.

Take an arbitrary $\tau \in [0, \infty)$ and subtract the equation

$$y_t^*(\tau) = y_t^*(0) + \int_0^\tau f(u_t^*(\tau'), y_t^*(\tau'), z^*(t)) d\tau' \quad (7.21)$$

from the equation

$$y_t^{N, M}(\tau) = y_t^*(0) + \int_0^\tau f(u^{N, M}(y_t^{N, M}(\tau'), z^*(t)), y_t^{N, M}(\tau'), z^*(t)) d\tau'. \quad (7.22)$$

We will obtain

$$\begin{aligned} \|y_t^{N, M}(\tau) - y_t^*(\tau)\| &\leq \int_0^\tau \|f(u^{N, M}(y_t^{N, M}(\tau'), z^*(t)), y_t^{N, M}(\tau'), z^*(t)) \\ &\quad - f(u_t^*(\tau'), y_t^*(\tau'), z^*(t))\| d\tau' \leq \int_0^\tau \|f(u^{N, M}(y_t^{N, M}(\tau'), z^*(t)), y_t^{N, M}(\tau'), z^*(t)) \\ &\quad - f(u^{N, M}(y_t^*(\tau'), z^*(t)), y_t^*(\tau'), z^*(t))\| d\tau' \\ &\quad + \int_0^\tau \|f(u^{N, M}(y_t^*(\tau'), z^*(t)), y_t^*(\tau'), z^*(t)) - f(u_t^*(\tau'), y_t^*(\tau'), z^*(t))\| d\tau'. \end{aligned} \quad (7.23)$$

Using Assumption 5.7 (ii),(iii), one can derive that

$$\int_0^\tau \|f(u^{N, M}(y_t^{N, M}(\tau'), z^*(t)), y_t^{N, M}(\tau'), z^*(t)) - f(u^{N, M}(y_t^*(\tau'), z^*(t)), y_t^*(\tau'), z^*(t))\| d\tau'$$

$$\begin{aligned}
&\leq \int_{\tau' \notin P_{t,\tau}(N,M)} \|f(u^{N,M}(y_t^{N,M}(\tau'), z^*(t)), y_t^{N,M}(\tau'), z^*(t)) - f(u^{N,M}(y_t^*(\tau'), z^*(t)), y_t^*(\tau'), z^*(t))\| d\tau' \\
&\quad + \int_{\tau' \in P_{t,\tau}(N,M)} [\|f(u^{N,M}(y_t^{N,M}(\tau'), z^*(t)), y_{z^*(t)}^{N,M}(\tau'), z^*(t))\| \\
&\quad + \|f(u^{N,M}(y_t^*(\tau'), z^*(t)), y_t^*(\tau'), z^*(t))\|] d\tau' \\
&\leq L_1 \int_0^\tau \|y_t^{N,M}(\tau') - y_t^*(\tau')\| d\tau' + L_2 \text{meas}\{P_{t,\tau}(N, M)\}, \tag{7.24}
\end{aligned}$$

where L_1 is a constant defined (in an obvious way) by Lipschitz constants of $f(\cdot)$ and $u^{N,M}(\cdot)$, and $L_2 \stackrel{\text{def}}{=} 2 \max_{(u,y,z) \in U \times Y \times Z} \{\|f(u, y, z)\|\}$.

Also, due to (7.19) and the dominated convergence theorem (see, e.g., p. 49 in [3])

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \int_0^\tau \|f(u^{N,M}(y_t^*(\tau'), z^*(t)), y_t^*(\tau'), z^*(t)) - f(u_t^*(\tau'), y_t^*(\tau'), z^*(t))\| d\tau' = 0. \tag{7.25}$$

Let us introduce the notation

$$\begin{aligned}
\kappa_{t,\tau}(N, M) &\stackrel{\text{def}}{=} L_2 \text{meas}\{P_{t,\tau}(N, M)\} \\
&+ \int_0^\tau \|f(u^{N,M}(y_t^*(\tau'), z^*(t)), y_t^*(\tau'), z^*(t)) - f(u_t^*(\tau'), y_t^*(\tau'), z^*(t))\| d\tau'
\end{aligned}$$

and rewrite the inequality (7.23) in the form

$$\|y_t^{N,M}(\tau) - y_t^*(\tau)\| \leq L_1 \int_0^\tau \|y_t^{N,M}(\tau') - y_t^*(\tau')\| d\tau' + \kappa_{t,\tau}(N, M). \tag{7.26}$$

By Gronwall-Bellman lemma, it follows that

$$\max_{\tau' \in [0, \tau]} \|y_t^{N,M}(\tau') - y_t^*(\tau')\| \leq \kappa_{t,\tau}(N, M) e^{L_1 \tau}. \tag{7.27}$$

Since, by (5.18) and (7.25),

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \kappa_{t,\tau}(N, M) = 0, \tag{7.28}$$

the inequality (7.27) implies (5.27).

By (5.27), $y_t^{N,M}(\tau) \in B_{t,\tau}$ for N and M large enough (for $\tau \in [0, \infty)$ such that the ball $B_{t,\tau}$ is not empty). Hence,

$$\begin{aligned}
\|u^{N,M}(y_t^{N,M}(\tau), z^*(t)) - u_t^*(\tau)\| &\leq \|u^{N,M}(y_t^{N,M}(\tau), z^*(t)) - u^{N,M}(y_t^*(\tau), z^*(t))\| \\
&+ \|u^{N,M}(y_t^*(\tau), z^*(t)) - u_t^*(\tau)\| \leq L \|y_t^{N,M}(\tau) - y_t^*(\tau)\| + \|u^{N,M}(y_t^*(\tau), z^*(t)) - u_t^*(\tau)\|.
\end{aligned}$$

The latter implies (5.28) (by (5.27) and (7.19)). \square

Proof of Theorem 5.8. Let $t \in [0, \infty)$ be such that Q_t is not empty and (5.5) is satisfied for an arbitrary $r > 0$. By (5.8) and (5.10), for any continuous $h(u, y)$ and for an arbitrary small $\alpha > 0$, there exists $S > 0$ such that

$$|S^{-1} \int_0^S h(u_t^*(\tau), y_t^*(\tau)) d\tau - \int_{U \times Y} h(u, y) \mu^*(t)(du, dy)| \leq \frac{\alpha}{2} \tag{7.29}$$

and

$$|S^{-1} \int_0^S h(u^{N,M}(y_t^{N,M}(\tau), z^*(t)), y_t^{N,M}(\tau)) d\tau - \int_{U \times Y} h(u, y) \mu^{N,M}(z^*(t))(du, dy)| \leq \frac{\alpha}{2}. \tag{7.30}$$

Using (7.30) and (7.29), one can obtain

$$\begin{aligned} & \left| \int_{U \times Y} h(u, y) \mu^{N, M}(z^*(t))(du, dy) - \int_{U \times Y} h(u, y) \mu^*(t)(du, dy) \right| \\ & \leq |S^{-1} \int_0^S h(u^{N, M}(y_t^{N, M}(\tau), z^*(t)), y_t^{N, M}(\tau)) d\tau - S^{-1} \int_0^S h(u_t^*(\tau), y_t^*(\tau)) d\tau| + \alpha. \end{aligned}$$

Due to Lemma 5.9, the latter implies the following inequality

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \left| \int_{U \times Y} h(u, y) \mu^{N, M}(z^*(t))(du, dy) - \int_{U \times Y} h(u, y) \mu^*(t)(du, dy) \right| \leq \alpha,$$

which, in turn, implies

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \left| \int_{U \times Y} h(u, y) \mu^{N, M}(z^*(t))(du, dy) - \int_{U \times Y} h(u, y) \mu^*(t)(du, dy) \right| = 0 \quad (7.31)$$

(due to the fact that α can be arbitrary small). Since $h(u, y)$ is an arbitrary continuous function, from (7.31) it follows that

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \rho(\mu^{N, M}(z^*(t)), \mu^*(t)) = 0. \quad (7.32)$$

Note that from the assumptions made it follows that Q_t is not empty and (5.5) is satisfied for an arbitrary $r > 0$ for almost all $t \in [0, \infty)$. Hence, (7.32) is valid for almost all $t \in [0, \infty)$.

Taking an arbitrary $t \in [0, \infty)$ and subtracting the equation

$$z^*(t) = z_0 + \int_0^t \tilde{g}(\mu^*(t'), z^*(t')) dt' \quad (7.33)$$

from the equation

$$z^{N, M}(t) = z_0 + \int_0^t \tilde{g}(\mu^{N, M}(z^{N, M}(t')), z^{N, M}(t')) dt', \quad (7.34)$$

one obtains

$$\begin{aligned} \|z^{N, M}(t) - z^*(t)\| & \leq \int_0^t \|\tilde{g}(\mu^{N, M}(z^{N, M}(t')), z^{N, M}(t')) - \tilde{g}(\mu^*(t'), z^*(t'))\| dt' \\ & \leq \int_0^t \|\tilde{g}(\mu^{N, M}(z^{N, M}(t')), z^{N, M}(t')) - \tilde{g}(\mu^{N, M}(z^*(t')), z^*(t'))\| dt' \\ & \quad + \int_0^t \|\tilde{g}(\mu^{N, M}(z^*(t')), z^*(t')) - \tilde{g}(\mu^*(t'), z^*(t'))\| dt'. \end{aligned} \quad (7.35)$$

From (5.13) and from the definition of the set $A_t(N, M)$ (see (5.16)), it follows that

$$\begin{aligned} & \int_0^t \|\tilde{g}(\mu^{N, M}(z^{N, M}(t')), z^{N, M}(t')) - \tilde{g}(\mu^{N, M}(z^*(t')), z^*(t'))\| dt' \\ & \leq \int_{t' \notin A_t(N, M)} \|\tilde{g}(\mu^{N, M}(z^{N, M}(t')), z^{N, M}(t')) - \tilde{g}(\mu^{N, M}(z^*(t')), z^*(t'))\| dt' \\ & \quad + \int_{t' \in A_t(N, M)} [\|\tilde{g}(\mu^{N, M}(z^{N, M}(t')), z^{N, M}(t'))\| + \|\tilde{g}(\mu^{N, M}(z^*(t')), z^*(t'))\|] dt' \\ & \leq L \int_0^t \|z^{N, M}(t') - z^*(t')\| + 2L_g \text{meas}\{A_t(N, M)\}, \end{aligned} \quad (7.36)$$

where $L_g \stackrel{\text{def}}{=} \max_{(u, y, z) \in U \times Y \times Z} \|g(u, y, z)\|$. This and (7.35) allows one to obtain the inequality

$$\|z^{N, M}(t) - z^*(t)\| \leq L \int_0^t \|z^{N, M}(t') - z^*(t')\| dt' + \kappa_t(N, M), \quad (7.37)$$

where

$$\kappa_t(N, M) \stackrel{\text{def}}{=} 2L_g \text{meas}\{A_t(N, M)\} + \int_0^t \|\tilde{g}(\mu^{N,M}(z^*(t')), z^*(t')) - \tilde{g}(\mu^*(t'), z^*(t'))\| dt'.$$

Note that, by (7.32),

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \int_0^t \|\tilde{g}(\mu^{N,M}(z^*(t')), z^*(t')) - \tilde{g}(\mu^*(t'), z^*(t'))\| dt' = 0, \quad (7.38)$$

which, along with (5.15), imply that

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \kappa_t(N, M) = 0. \quad (7.39)$$

By Gronwall-Bellman lemma, from (7.37) it follows that

$$\max_{t' \in [0, t]} \|z^{N,M}(t') - z^*(t')\| \leq \kappa_t(N, M)e^{Lt}.$$

The latter along with (7.39) imply (5.22).

Let us now establish the validity of (5.23). Let $t \in [0, \infty)$ be such that the ball Q_t introduced in Assumption 5.6 is not empty. By triangle inequality,

$$\rho(\mu^{N,M}(z^{N,M}(t)), \mu^*(t)) \leq \rho(\mu^{N,M}(z^{N,M}(t)), \mu^{N,M}(z^*(t))) + \rho(\mu^{N,M}(z^*(t)), \mu^*(t)) \quad (7.40)$$

Due to (5.22), $z^{N,M}(t) \in Q_t$ for M and N large enough. Hence, by (5.12),

$$\rho(\mu^{N,M}(z^{N,M}(t)), \mu^{N,M}(z^*(t))) \leq \kappa(\|z^{N,M}(t) - z^*(t)\|),$$

which implies that

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \rho(\mu^{N,M}(z^{N,M}(t)), \mu^{N,M}(z^*(t))) = 0.$$

The latter, along with (7.32) and (7.40), imply (5.23).

Finally, let us prove (5.24). By (5.6) and (5.11), for any continuous function $\tilde{h}(\mu, z) : \mathcal{P}(\mathcal{P}(U \times Y) \times Z) \rightarrow \mathbb{R}^1$, and for an arbitrary small $\alpha > 0$, there exists $\tilde{T} > 0$ such that

$$|\tilde{T}^{-1} \int_0^{\tilde{T}} \tilde{h}(\mu^*(t), z^*(t)) dt - \int_F \tilde{h}(\mu, z) p^*(d\mu, dz)| \leq \frac{\alpha}{2} \quad (7.41)$$

and

$$|\tilde{T}^{-1} \int_0^{\tilde{T}} \tilde{h}(\mu^{N,M}(t), z^{N,M}(t)) dt - \int_F \tilde{h}(\mu, z) \lambda^{N,M}(d\mu, dz)| \leq \frac{\alpha}{2}. \quad (7.42)$$

Using (7.42) and (7.41), one can obtain

$$\begin{aligned} & \left| \int_F \tilde{h}(\mu, z) \lambda^{N,M}(d\mu, dz) - \int_F \tilde{h}(\mu, z) p^*(d\mu, dz) \right| \\ & \leq |\tilde{T}^{-1} \int_0^{\tilde{T}} \tilde{h}(\mu^{N,M}(t), z^{N,M}(t)) dt - \tilde{T}^{-1} \int_0^{\tilde{T}} \tilde{h}(\mu^*(t), z^*(t)) dt| + \alpha. \end{aligned}$$

Due to (5.22) and (5.23), the latter implies the following inequality

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \left| \int_F \tilde{h}(\mu, z) \lambda^{N,M}(d\mu, dz) - \int_F \tilde{h}(\mu, z) p^*(d\mu, dz) \right| \leq \alpha,$$

which, in turn, implies

$$\lim_{N \rightarrow \infty} \limsup_{M \rightarrow \infty} \left| \int_F \tilde{h}(\mu, z) \lambda^{N,M}(d\mu, dz) - \int_F \tilde{h}(\mu, z) p^*(d\mu, dz) \right| = 0 \quad (7.43)$$

(due to the fact that α can be arbitrary small). This proves (5.21). Taking now $\tilde{h}(\mu, z) = \tilde{G}(\mu, z)$ in (7.43) and having in mind that

$$\int_F \tilde{G}(\mu, z) \lambda^{N,M}(d\mu, dz) = \lim_{\mathcal{T} \rightarrow \infty} \mathcal{T}^{-1} \int_0^{\mathcal{T}} \tilde{G}(\mu^{N,M}(t), z^{N,M}(t)) dt$$

(see (5.11)) and that

$$\int_F \tilde{G}(\mu, z) p^*(d\mu, dz) = \tilde{G}^*,$$

one proves the validity of (5.24). This completes the proof of the theorem. \square

8. Proof of Theorem 6.3. Denote by $z(t)$ the solution of the differential equation

$$z'(t) = \tilde{g}(\mu^{N,M}(t), z(t)) \quad (8.1)$$

considered on the interval $[\mathcal{T}_0, \mathcal{T}_0 + \mathcal{T}]$ that satisfies the initial condition

$$z(\mathcal{T}_0) = z \in Z. \quad (8.2)$$

Also, denote by $\bar{z}(t)$ the solution of the differential equation

$$z'(t) = \tilde{g}(\bar{\mu}^{N,M}(t), z(t)) \quad (8.3)$$

considered on the same interval $[\mathcal{T}_0, \mathcal{T}_0 + \mathcal{T}]$ and satisfying the same initial condition (8.2), where $\bar{\mu}^{N,M}(t)$ is the piecewise constant function defined as follows

$$\bar{\mu}^{N,M}(t) \stackrel{\text{def}}{=} \mu^{N,M}(t_l) \quad \forall t \in [t_l, t_{l+1}), \quad l = 0, 1, \dots \quad (8.4)$$

Using the piecewise continuity property (6.14), it can be readily established (using a standard argument, see, e.g., the proof of Theorem 4.5 in [48]) that

$$\max_{t \in [\mathcal{T}_0, \mathcal{T}_0 + \mathcal{T}]} \|\bar{z}(t) - z(t)\| \leq \kappa_1(\epsilon, \mathcal{T}), \quad \text{where} \quad \lim_{\epsilon \rightarrow 0} \kappa_1(\epsilon, \mathcal{T}) = 0. \quad (8.5)$$

The latter implies, in particular,

$$\max_{t \in [0, \mathcal{T}_0]} \|z^{N,M}(t) - \bar{z}^{N,M}(t)\| \leq \kappa_1(\epsilon, \mathcal{T}_0), \quad (8.6)$$

where $\bar{z}^{N,M}(t)$ is the solution of (8.3) that satisfies the initial condition $\bar{z}^{N,M}(0) = z^{N,M}(0)$.

Choose now \mathcal{T}_0 in such a way that

$$c_3 e^{-c_4 \mathcal{T}_0} \stackrel{\text{def}}{=} a < 1 \quad (8.7)$$

and denote by $\bar{z}_1^{N,M}(t)$ the solution of the system (8.3) considered on the interval $[\mathcal{T}_0, 2\mathcal{T}_0]$ with the initial condition $\bar{z}_1^{N,M}(\mathcal{T}_0) = z^{N,M}(\mathcal{T}_0)$. From (6.5) and (8.6) it follows that

$$\|\bar{z}_1^{N,M}(2\mathcal{T}_0) - \bar{z}^{N,M}(2\mathcal{T}_0)\| \leq a \|z^{N,M}(\mathcal{T}_0) - \bar{z}^{N,M}(\mathcal{T}_0)\| \leq a \kappa_1(\epsilon, \mathcal{T}_0). \quad (8.8)$$

Also, taking into account the validity of (8.5), one can write down

$$\begin{aligned} \|z^{N,M}(2\mathcal{T}_0) - \bar{z}^{N,M}(2\mathcal{T}_0)\| &\leq \|z^{N,M}(2\mathcal{T}_0) - \bar{z}_1^{N,M}(2\mathcal{T}_0)\| + \|\bar{z}_1^{N,M}(2\mathcal{T}_0) - \bar{z}^{N,M}(2\mathcal{T}_0)\| \\ &\leq \kappa_1(\epsilon, \mathcal{T}_0) + a \kappa_1(\epsilon, \mathcal{T}_0) \leq \frac{\kappa_1(\epsilon, \mathcal{T}_0)}{1-a}. \end{aligned}$$

By continuing in a similar way, one can prove that, for any $k = 1, 2, \dots$,

$$\|z^{N,M}(k\mathcal{T}_0) - \bar{z}^{N,M}(k\mathcal{T}_0)\| \leq (1 + a + \dots + a^{k-1}) \kappa_1(\epsilon, \mathcal{T}_0) \leq \frac{\kappa_1(\epsilon, \mathcal{T}_0)}{1-a}.$$

Hence, by (6.5),

$$\max_{t \in [k\mathcal{T}_0, (k+1)\mathcal{T}_0]} \|z^{N,M}(t) - \bar{z}^{N,M}(t)\| \leq c_3 \frac{\kappa_1(\epsilon, \mathcal{T}_0)}{1-a} \quad \forall k = 0, 1, \dots$$

and, consequently,

$$\sup_{t \in [0, \infty)} \|z^{N,M}(t) - \bar{z}^{N,M}(t)\| \leq c_3 \frac{\kappa_1(\epsilon, \mathcal{T}_0)}{1-a} \stackrel{\text{def}}{=} \kappa_2(\epsilon), \quad \lim_{\epsilon \rightarrow 0} \kappa_2(\epsilon) = 0. \quad (8.9)$$

Using (8.9), one can obtain

$$\begin{aligned} & \left| \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{G}(\mu^{N,M}(t), z^{N,M}(t)) dt - \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{G}(\bar{\mu}^{N,M}(t), \bar{z}^{N,M}(t)) dt \right| \\ & \leq \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} |\tilde{G}(\mu^{N,M}(t), z^{N,M}(t)) - \tilde{G}(\bar{\mu}^{N,M}(t), \bar{z}^{N,M}(t))| dt + L\kappa_2(\epsilon) \\ & \leq \frac{1}{\mathcal{T}} \sum_{l=0}^{\lfloor \frac{\mathcal{T}}{\Delta(\epsilon)} \rfloor} \int_{t_l}^{t_{l+1}} |\tilde{G}(\mu^{N,M}(t), z^{N,M}(t)) - \tilde{G}(\mu^{N,M}(t_l), z^{N,M}(t_l))| dt + L\kappa_2(\epsilon) + 2M\Delta(\epsilon) \quad \forall \mathcal{T} \geq 1, \end{aligned} \quad (8.10)$$

where $\lfloor \cdot \rfloor$ stands for the floor function ($\lfloor x \rfloor$ is the maximal integer number that is less or equal than x), L is a Lipschitz constant (for simplicity it is assumed that $G(u, y, z)$ is Lipschitz continuous in z) and $M \stackrel{\text{def}}{=} \max_{(u, y, z) \in U \times Y \times Z} |G(u, y, z)|$.

Without loss of generality, one may assume that r_δ is decreasing with δ and that $r_\delta \leq \delta$ (the later can be achieved by replacing r_δ with $\min\{\delta, r_\delta\}$ if necessary). Having this in mind, define $\delta(\epsilon)$ as the solution of the problem

$$\min\{\delta : r_\delta \geq \Delta^{\frac{1}{2}}(\epsilon)\}. \quad (8.11)$$

That is,

$$r_{\delta(\epsilon)} = \Delta^{\frac{1}{2}}(\epsilon). \quad (8.12)$$

Note that, by construction,

$$\lim_{\epsilon \rightarrow 0} \delta(\epsilon) = 0, \quad \delta(\epsilon) \geq \Delta^{\frac{1}{2}}(\epsilon). \quad (8.13)$$

By (6.14),

$$\int_{t_l}^{t_{l+1}} |\tilde{G}(\mu^{N,M}(t), z^{N,M}(t)) - \tilde{G}(\mu^{N,M}(t_l), z^{N,M}(t_l))| dt \leq \Delta(\epsilon) \nu(\Delta(\epsilon)) \quad \text{if } t_l \notin \cup_{i=1}^k (\mathcal{T}_i - \delta(\epsilon), \mathcal{T}_i + \delta(\epsilon)). \quad (8.14)$$

Also,

$$\int_{t_l}^{t_{l+1}} |\tilde{G}(\mu^{N,M}(t), z^{N,M}(t)) - \tilde{G}(\mu^{N,M}(t_l), z^{N,M}(t_l))| dt \leq 2M\Delta(\epsilon) \quad \text{if } t_l \in \cup_{i=1}^k (\mathcal{T}_i - \delta(\epsilon), \mathcal{T}_i + \delta(\epsilon)). \quad (8.15)$$

Taking (6.13), (8.14) and (8.15) into account, one can use (8.10) to obtain the following estimate

$$\begin{aligned} & \left| \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{G}(\mu^{N,M}(t), z^{N,M}(t)) dt - \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{G}(\bar{\mu}^{N,M}(t), \bar{z}^{N,M}(t)) dt \right| \\ & \leq \frac{1}{\mathcal{T}} \lfloor \frac{\mathcal{T}}{\Delta(\epsilon)} \rfloor \Delta(\epsilon) \nu(\Delta(\epsilon)) + \frac{1}{\mathcal{T}} (c\mathcal{T}) \lfloor \frac{2\delta(\epsilon)}{\Delta(\epsilon)} + 2 \rfloor (2M\Delta(\epsilon)) + L\kappa_2(\epsilon) + 2M\Delta(\epsilon) \stackrel{\text{def}}{=} \kappa_3(\epsilon) \quad \forall \mathcal{T} \geq 1. \end{aligned}$$

Thus,

$$\sup_{\mathcal{T} \geq 1} \left| \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{G}(\mu^{N,M}(t), z^{N,M}(t)) dt - \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{G}(\bar{\mu}^{N,M}(t), \bar{z}^{N,M}(t)) dt \right| \leq \kappa_3(\epsilon), \quad \lim_{\epsilon \rightarrow 0} \kappa_3(\epsilon) = 0. \quad (8.16)$$

Denote by $\bar{z}(t)$ the solution of the differential equation (8.3) considered on the interval $[\mathcal{T}_0, \mathcal{T}_0 + \mathcal{T}]$ and satisfying the initial condition $\bar{z}(\mathcal{T}_0) = z_\epsilon^{N,M}(\mathcal{T}_0)$, where $\mathcal{T}_0 \stackrel{\text{def}}{=} l_0 \Delta(\epsilon)$ for some $l_0 \geq 0$. Subtracting the equation

$$\bar{z}(t_{l+1}) = \bar{z}(t_l) + \int_{t_l}^{t_{l+1}} \tilde{g}(\mu^{N,M}(t_l), \bar{z}(t)) dt, \quad l \geq l_0,$$

from the equation

$$z_\epsilon^{N,M}(t_{l+1}) = z_\epsilon^{N,M}(t_l) + \int_{t_l}^{t_{l+1}} g(u_{z^{N,M}(t_l)}(\frac{t-t_l}{\epsilon}), y_\epsilon^{N,M}(t), z_\epsilon^{N,M}(t)) dt, \quad l \geq l_0,$$

one can obtain

$$\begin{aligned} & \|z_\epsilon^{N,M}(t_{l+1}) - \bar{z}(t_{l+1})\| \leq \|z_\epsilon^{N,M}(t_l) - \bar{z}(t_l)\| \\ & + \int_{t_l}^{t_{l+1}} \|g(u_{z^{N,M}(t_l)}(\frac{t-t_l}{\epsilon}), y_\epsilon^{N,M}(t), z_\epsilon^{N,M}(t)) - g(u_{z^{N,M}(t_l)}(\frac{t-t_l}{\epsilon}), y_\epsilon^{N,M}(t), z_\epsilon^{N,M}(t_l))\| dt \\ & + \|\int_{t_l}^{t_{l+1}} g(u_{z^{N,M}(t_l)}(\frac{t-t_l}{\epsilon}), y_\epsilon^{N,M}(t), z_\epsilon^{N,M}(t_l)) dt - \Delta(\epsilon) \tilde{g}(\mu^{N,M}(t_l), z_\epsilon^{N,M}(t_l))\| \\ & + \int_{t_l}^{t_{l+1}} \|\tilde{g}(\mu^{N,M}(t_l), z_\epsilon^{N,M}(t_l)) - \tilde{g}(\mu^{N,M}(t_l), \bar{z}(t))\| dt \\ & \leq \|z_\epsilon^{N,M}(t_l) - \bar{z}(t_l)\| + L_1 \Delta(\epsilon) \|z_\epsilon^{N,M}(t_l) - \bar{z}(t_l)\| + L_2 \Delta^2(\epsilon) + \Delta(\epsilon) \bar{\phi}_g(\frac{\Delta(\epsilon)}{\epsilon}), \end{aligned} \quad (8.17)$$

where L_i , $i = 1, 2$, are positive constants and $\bar{\phi}_g(\cdot)$ is defined in (6.7). Note that, in order to obtain the estimate above, one needs to take into account the fact that

$$\max\left\{\max_{t \in [t_l, t_{l+1}]} \{\|z_\epsilon^{N,M}(t) - z_\epsilon^{N,M}(t_l)\|\}, \max_{t \in [t_l, t_{l+1}]} \{\|\bar{z}(t) - \bar{z}(t_l)\|\}\right\} \leq L_3 \Delta(\epsilon), \quad L_3 > 0 \quad (8.18)$$

as well as the fact that (see (6.7))

$$\begin{aligned} & \|\int_{t_l}^{t_{l+1}} g(u_{z^{N,M}(t_l)}(\frac{t-t_l}{\epsilon}), y_\epsilon^{N,M}(t), z_\epsilon^{N,M}(t_l)) dt - \Delta(\epsilon) \tilde{g}(\mu^{N,M}(t_l), z_\epsilon^{N,M}(t_l))\| \\ & = \Delta(\epsilon) \left[\left(\frac{\Delta(\epsilon)}{\epsilon}\right)^{-1} \int_0^{\frac{\Delta(\epsilon)}{\epsilon}} g(u_{z^{N,M}(t_l)}(\tau), y_{z^{N,M}(t_l)}^{N,M}(\tau, y_\epsilon^{N,M}(t_l)), z_\epsilon^{N,M}(t_l)) d\tau - \tilde{g}(\mu^{N,M}(t_l), z_\epsilon^{N,M}(t_l)) \right] \\ & \leq \Delta(\epsilon) \bar{\phi}_g(\frac{\Delta(\epsilon)}{\epsilon}), \end{aligned} \quad (8.19)$$

where $\tau = \frac{t-t_l}{\epsilon}$ and $y_{z^{N,M}(t_l)}^{N,M}(\tau, y_\epsilon^{N,M}(t_l)) = y_\epsilon^{N,M}(t_l + \epsilon\tau)$. From (8.17) it follows (see Proposition 5.1 in [39])

$$\|z_\epsilon^{N,M}(t_l) - \bar{z}(t_l)\| \leq \kappa_4(\epsilon, \mathcal{T}), \quad l = l_0, l_0 + 1, \dots, l_0 + \lfloor \frac{\mathcal{T}}{\Delta(\epsilon)} \rfloor, \quad \lim_{\epsilon \rightarrow 0} \kappa_4(\epsilon) = 0.$$

This (due to (8.18)) leads to

$$\max_{t \in [\mathcal{T}_0, \mathcal{T}_0 + \mathcal{T}]} \|z_\epsilon^{N,M}(t) - \bar{z}(t)\| \leq \kappa_5(\epsilon, \mathcal{T}), \quad \text{where} \quad \lim_{\epsilon \rightarrow 0} \kappa_5(\epsilon, \mathcal{T}) = 0, \quad (8.20)$$

and, in particular, to

$$\max_{t \in [0, \mathcal{T}_0]} \|z_\epsilon^{N,M}(t) - \bar{z}^{N,M}(t)\| \leq \kappa_5(\epsilon, \mathcal{T}_0) \quad (8.21)$$

(since, by definition, $z_\epsilon^{N,M}(0) = z^{N,M}(0)$ and $\bar{z}^{N,M}(0) = z^{N,M}(0)$). Assume that \mathcal{T}_0 is chosen in such a way that (8.7) is satisfied and denote by $\bar{z}_1(t)$ the solution of the system (8.3) considered on the interval $[\mathcal{T}_0, 2\mathcal{T}_0]$ with the initial condition $\bar{z}_1(\mathcal{T}_0) = z_\epsilon^{N,M}(\mathcal{T}_0)$. By (6.5) and (8.21),

$$\|\bar{z}_1(2\mathcal{T}_0) - \bar{z}^{N,M}(2\mathcal{T}_0)\| \leq a \|z_\epsilon^{N,M}(\mathcal{T}_0) - \bar{z}^{N,M}(\mathcal{T}_0)\| \leq a \kappa_5(\epsilon, \mathcal{T}_0). \quad (8.22)$$

Also, by (8.20),

$$\begin{aligned} \|z_\epsilon^{N,M}(2\mathcal{T}_0) - \bar{z}^{N,M}(2\mathcal{T}_0)\| &\leq \|z_\epsilon^{N,M}(2\mathcal{T}_0) - \bar{z}_1(2\mathcal{T}_0)\| + \|\bar{z}_1(2\mathcal{T}_0) - \bar{z}^{N,M}(2\mathcal{T}_0)\| \\ &\leq \kappa_5(\epsilon, \mathcal{T}_0) + a\kappa_5(\epsilon, \mathcal{T}_0) \leq \frac{\kappa_5(\epsilon, \mathcal{T}_0)}{1-a}. \end{aligned}$$

Continuing in a similar way, one can prove that, for any $k = 1, 2, \dots$,

$$\|z_\epsilon^{N,M}(k\mathcal{T}_0) - \bar{z}^{N,M}(k\mathcal{T}_0)\| \leq (1 + a + \dots + a^{k-1})\kappa_5(\epsilon, \mathcal{T}_0) \leq \frac{\kappa_5(\epsilon, \mathcal{T}_0)}{1-a}. \quad (8.23)$$

Denote by $\bar{z}_k(t)$ the solution of the system (8.3) considered on the interval $[k\mathcal{T}_0, (k+1)\mathcal{T}_0]$ with the initial condition $\bar{z}_k(k\mathcal{T}_0) = z_\epsilon^{N,M}(k\mathcal{T}_0)$. By (6.5) and (8.21),

$$\begin{aligned} \max_{t \in [k\mathcal{T}_0, (k+1)\mathcal{T}_0]} \|z_\epsilon^{N,M}(t) - \bar{z}^{N,M}(t)\| &\leq \max_{t \in [k\mathcal{T}_0, (k+1)\mathcal{T}_0]} \|z_\epsilon^{N,M}(t) - \bar{z}_k(t)\| + \max_{t \in [k\mathcal{T}_0, (k+1)\mathcal{T}_0]} \|\bar{z}_k(t) - \bar{z}^{N,M}(t)\| \\ &\leq \kappa_5(\epsilon, \mathcal{T}_0) + c_3 \|z_\epsilon^{N,M}(k\mathcal{T}_0) - \bar{z}^{N,M}(k\mathcal{T}_0)\|. \end{aligned}$$

Thus, by (8.23),

$$\sup_{t \in [0, \infty)} \|z_\epsilon^{N,M}(t) - \bar{z}^{N,M}(t)\| \leq \kappa_5(\epsilon, \mathcal{T}_0) + c_3 \frac{\kappa_5(\epsilon, \mathcal{T}_0)}{1-a} \stackrel{\text{def}}{=} \kappa_6(\epsilon), \quad \lim_{\epsilon \rightarrow 0} \kappa_6(\epsilon) = 0. \quad (8.24)$$

From (8.9) and (8.24) it also follows that

$$\begin{aligned} &|\frac{1}{\mathcal{T}} \int_0^\mathcal{T} G(u_\epsilon^{N,M}(t), y_\epsilon^{N,M}(t), z_\epsilon^{N,M}(t))dt - \frac{1}{\mathcal{T}} \int_0^\mathcal{T} \tilde{G}(\bar{\mu}^{N,M}(t), \bar{z}^{N,M}(t))dt| \\ &\leq |\frac{1}{\mathcal{T}} \int_0^\mathcal{T} G(u_\epsilon^{N,M}(t), y_\epsilon^{N,M}(t), z^{N,M}(t))dt - \frac{1}{\mathcal{T}} \int_0^\mathcal{T} \tilde{G}(\bar{\mu}^{N,M}(t), z^{N,M}(t))dt| + L(\kappa_2(\epsilon) + \kappa_6(\epsilon)) \\ &\leq \frac{1}{\mathcal{T}} \sum_{l=0}^{\lfloor \frac{\mathcal{T}}{\Delta(\epsilon)} \rfloor} |\int_{t_l}^{t_{l+1}} G(u_{z^{N,M}(t_l)}^{N,M}(\frac{t-t_l}{\epsilon}), y_\epsilon^{N,M}(t), z^{N,M}(t_l))dt - \int_{t_l}^{t_{l+1}} \tilde{G}(\mu^{N,M}(t_l), z^{N,M}(t_l))dt| \\ &\quad + L(\kappa_2(\epsilon) + \kappa_6(\epsilon)) + 2M\Delta(\epsilon) + 2LL_3\Delta(\epsilon) \quad \forall \mathcal{T} \geq 1, \end{aligned} \quad (8.25)$$

where L and M are as in (8.10) and it has been taking into account that $\max_{t \in [t_l, t_{l+1}]} \|z^{N,M}(t) - z^{N,M}(t_l)\| \leq L_3\Delta(\epsilon)$, with L_3 being the same constant as in (8.18). Similarly to (8.19), one can obtain (using (6.7))

$$\begin{aligned} &|\int_{t_l}^{t_{l+1}} G(u_{z^{N,M}(t_l)}^{N,M}(\frac{t-t_l}{\epsilon}), y_\epsilon^{N,M}(t), z^{N,M}(t_l))dt - \Delta(\epsilon)\tilde{G}(\mu^{N,M}(t_l), z^{N,M}(t_l))| \\ &= \Delta(\epsilon) [(\frac{\Delta(\epsilon)}{\epsilon})^{-1} \int_0^{\frac{\Delta(\epsilon)}{\epsilon}} G(u_{z^{N,M}(t_l)}^{N,M}(\tau), y_{z^{N,M}(t_l)}^{N,M}(\tau), y_\epsilon^{N,M}(t_l), z^{N,M}(t_l))d\tau - \tilde{G}(\mu^{N,M}(t_l), z^{N,M}(t_l))] \\ &\leq \Delta(\epsilon)\bar{\phi}_G(\frac{\Delta(\epsilon)}{\epsilon}), \end{aligned}$$

The latter along with (8.25) imply that

$$|\frac{1}{\mathcal{T}} \int_0^\mathcal{T} G(u_\epsilon^{N,M}(t), y_\epsilon^{N,M}(t), z_\epsilon^{N,M}(t))dt - \frac{1}{\mathcal{T}} \int_0^\mathcal{T} \tilde{G}(\bar{\mu}^{N,M}(t), \bar{z}^{N,M}(t))dt| \leq \kappa_7(\epsilon) \quad \forall \mathcal{T} \geq 1, \quad (8.26)$$

where

$$\kappa_7(\epsilon) \stackrel{\text{def}}{=} L(\kappa_2(\epsilon) + \kappa_6(\epsilon)) + 2M\Delta(\epsilon) + 2LL_3\Delta(\epsilon) + \bar{\phi}_G(\frac{\Delta(\epsilon)}{\epsilon}), \quad \lim_{\epsilon \rightarrow 0} \kappa_7(\epsilon) = 0.$$

Hence, by (8.16),

$$|\frac{1}{\mathcal{T}} \int_0^\mathcal{T} G(u_\epsilon^{N,M}(t), y_\epsilon^{N,M}(t), z_\epsilon^{N,M}(t))dt - \frac{1}{\mathcal{T}} \int_0^\mathcal{T} \tilde{G}(\mu^{N,M}(t), z^{N,M}(t))dt| \leq \kappa_3(\epsilon) + \kappa_7(\epsilon) \quad \forall \mathcal{T} \geq 1, \quad (8.27)$$

and, consequently,

$$\begin{aligned} & \left| \liminf_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} G(u_\epsilon^{N,M}(t), y_\epsilon^{N,M}(t), z_\epsilon^{N,M}(t)) dt - \tilde{V}^{N,M} \right| \leq \kappa_3(\epsilon) + \kappa_7(\epsilon), \\ \Rightarrow & \lim_{\epsilon \rightarrow 0} \liminf_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} G(u_\epsilon^{N,M}(t), y_\epsilon^{N,M}(t), z_\epsilon^{N,M}(t)) dt = \tilde{V}^{N,M} = \tilde{G}^* + \beta(N, M) \end{aligned}$$

(see (5.25) and (5.26)). Due to (2.34), the latter proves the $\beta(N, M)$ -asymptotic near optimality of the control $u_\epsilon^{N,M}(\cdot)$. Also, the estimate (6.15) follows from (8.9) and (8.24).

Using an arbitrary Lipschitz continuous function $h(u, y, z)$ instead of $G(u, y, z)$, one can obtain (similarly to (8.27))

$$\left| \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} h(u_\epsilon^{N,M}(t), y_\epsilon^{N,M}(t), z_\epsilon^{N,M}(t)) dt - \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \tilde{h}(\mu^{N,M}(t), z^{N,M}(t)) dt \right| \leq \kappa_8(\epsilon) \quad \forall \mathcal{T} \geq 1, \quad (8.28)$$

where $\lim_{\epsilon \rightarrow \infty} \kappa_8(\epsilon) = 0$. If the triplets $(u_\epsilon^{N,M}(\cdot), y_\epsilon^{N,M}(\cdot), z_\epsilon^{N,M}(\cdot))$ generates the occupational measure $\gamma_\epsilon^{N,M}$, then (see (1.5))

$$\lim_{\mathcal{T} \rightarrow \infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} h(u_\epsilon^{N,M}(t), y_\epsilon^{N,M}(t), z_\epsilon^{N,M}(t)) dt = \int_{U \times Y \times Z} h(u, y, z) \gamma_\epsilon^{N,M}(du, dy, dz).$$

Hence, passing to the limit in (8.28) with $\mathcal{T} \rightarrow \infty$ and taking into account (5.11), one obtains

$$\left| \int_{U \times Y \times Z} h(u, y, z) \gamma_\epsilon^{N,M}(du, dy, dz) - \int_F \tilde{h}(\mu, z) \lambda^{N,M}(d\mu, dz) \right| \leq \kappa_8(\epsilon).$$

By the definition of the map $\Phi(\cdot)$ (see (2.40)), the latter implies that

$$\left| \int_{U \times Y \times Z} h(u, y, z) \gamma_\epsilon^{N,M}(du, dy, dz) - \int_{U \times Y \times Z} h(u, y, z) \Phi(\lambda^{N,M})(du, dy, dz) \right| \leq \kappa_8(\epsilon),$$

which, in turn, implies (6.16). This completes the proof. \square

REFERENCES

- [1] O. Alvarez and M. Bardi, “Viscosity Solutions Methods for Singular Perturbations in Deterministic and Stochastic Control”, *SIAM J. Control and Optimization*, 40 (2001): 4, pp.1159-1188.
- [2] O. Alvarez and M. Bardi, “Singular Perturbations of Nonlinear Degenerate Parabolic PDEs: a General Convergence result”, *Arch. Ration. Mech. Anal.*, 170 (2003), pp. 17-61.
- [3] R. B. Ash, “Real Analysis and Probability”, Academic Press, New York, 1972
- [4] E.J. Anderson and P. Nash, “Linear Programming in Infinite-Dimensional Spaces”, Wiley, Chichester, 1987.
- [5] A. Arapostathis, V.S. Borkar, M.K Ghosh, “Ergodic Control of Diffusion Processes”, Cambridge University Press, Cambridge, 2012.
- [6] Z. Artstein, “Invariant Measures of Differential Inclusions Applied to Singular Perturbations”, *J. Differential Equations*, 152 (1999), pp. 289-307.
- [7] Z. Artstein, “The Chattering Limit of Singularly Perturbed Optimal Control Problems”, *Proceedings of CDC-2000, Control and Decision Conference*, Sydney, 2000, pp. 564-569.
- [8] Z. Artstein, “Invariant Measures and Their Projections in Nonautonomous Dynamical Systems”, *Stochastics and Dynamics*, 4 (2004): 3, pp 439-459.
- [9] Z. Artstein and I. Bright, “Periodic Optimization Suffices for Infinite Horizon Planar Optimal Control”, *SIAM J. Control and Optimization*, 48 (2010): 8, pp. 4963-4986
- [10] Z. Artstein and V. Gaitsgory, “Tracking Fast Trajectories Along a Slow Dynamics: A Singular Perturbations Approach”, *SIAM J. Control and Optimization*, 35 (1997): 5, pp. 1487-1507.
- [11] Z. Artstein and V. Gaitsgory, “The value function of singularly perturbed control systems”, *Applied Mathematics and Optimization* 41(2000): 3, pp. 425-445
- [12] K. Avrachenkov, J. Filar and M. Haviv, “Singular perturbations of Markov chains and decision processes. A Survey”, A chapter in “Handbook of Markov Decision Processes: Methods and Applications”, E.A. Feinberg, A. Shwartz (eds), *International Series in Operations Research and Management Science*, 40 (2002), pp.113-153, Kluwer Academic Publishers.
- [13] J.-P. Aubin, “Viability Theory”, Birkhauser, Boston, 1991.
- [14] M. Bardi and I. Capuzzo-Dolcetta, “Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations”, Birkhauser, Boston, 1997.

- [15] G.K. Basak, V.S. Borkar and M.K. Ghosh, "Ergodic Control of Degenerate Diffusions, *Stochastic Analysis and Applications*, 15 (1997), pp. 1-17.
- [16] A.G. Bhatt and V.S. Borkar, *Occupation Measures for Controlled Markov Processes: Characterization and Optimality*, Annals of Probability 24 (1996), 1531-1562.
- [17] A. Bensoussan, "Perturbation Methods in Optimal Control", John Wiley, New York, 1989.
- [18] V.S. Borkar and V. Gaitsgory, "On Existence of Limit Occupational Measures Set of a Controlled Stochastic Differential Equation, *SIAM J. on Control and Optimization*, 44 (2005/2006): 4, pp. 1436-1473.
- [19] V.S. Borkar and V. Gaitsgory, "On Averaging of Singularly Perturbed Controlled Stochastic Differential Equations", *Applied Mathematics and Optimization*, 56 (2007), pp. 169-209.
- [20] V.S. Borkar and V. Gaitsgory, "Singular Perturbations in Ergodic Control of Diffusions, *SIAM J. on Control and Optimization*, 46 (2007): 5, pp. 1562-1577.
- [21] R. Buckdahn, D. Goreac, and M. Quincampoix, "Stochastic optimal control and linear programming approach", *Appl. Math. Optim.* 63 (2011): 2, pp. 257-276.
- [22] F. Colonius, "Optimal Periodic Control", Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1988.
- [23] F. Colonius and R. Fabri, "Controllability for Systems with Slowly Varying Parameters", *ESAIM: Control, Optimisation and Calculus of Variations*, 9(2003), pp. 207-216.
- [24] F. Colonius and W. Kliemann, "Infinite Time Optimal Control and Periodicity", *Applied Mathematics and Optimization*, 20 (1989), pp 113-130.
- [25] G.B. Dantzig, "Linear Programming and Extensions", Princeton University Press, Princeton, 1963.
- [26] M.G. Dmitriev and G.A. Kurina, "Singular Perturbations in Control Problems (Review)", *Automation and Remote Control*, 67 (2006): 1, pp. 1-43
- [27] T.D. Donchev, "Singularly Perturbed Evolution Inclusions", *SIAM J. Control and Optimization*, 48(2010): 7, pp. 4572-4590.
- [28] T.D. Donchev and A.L. Dontchev, "Singular Perturbations in Infinite-Dimensional Control Systems", *SIAM J. Control and Optimization*, 42 (2003), pp. 1795-1812.
- [29] A. Dontchev, T. Donchev and I. Slavov, "A Tichonov -type Theorem for Singularly Perturbed Differential Inclusion", *Nonlinear Analysis*, 26 (1996), 1547-1554.
- [30] T. Dontchev and I. Slavov, "Averaging Method for One Sided Lipschitz Differential Inclusions", *SIAM J. Control and Optimization*, 37 (1999), pp 1600-1613.
- [31] L.C. Evans and D. Gomes, "Linear Programming Interpretations of Mather's Variational Principle", *Optimization and Calculus of Variations* 8 (2002).
- [32] M. Falcone, "Numerical Solution of Dynamic Programming Equations", Appendix A in [14].
- [33] J.A. Filar, V. Gaitsgory and A. Haurie, "Control of Singularly Perturbed Hybrid Stochastic Systems", *IEEE Trans. on Automatic Control*, 46 (2001): 2.
- [34] O.P. Filatov and M.M. Hapaev, "Averaging of Systems of Differential Inclusions", Moscow University Publishing House, Moscow, 1998 (in Russian).
- [35] L. Finlay, V. Gaitsgory and I. Lebedev, "Duality In Linear Programming Problems Related to Long Run Average Problems of Optimal Control", *SIAM J. on Control and Optimization*, 47 (2008): 4, pp. 1667-1700.
- [36] W.H. Fleming and H.M. Soner, "Controlled Markov Processes and Viscosity Solutions", Springer-Verlag, New York, 1991.
- [37] W. H. Fleming and D. Vermes, "Convex duality approach to the optimal control of diffusions", *SIAM J. Control Optimization*, 27 (1989): 5, pp. 1136-1155.
- [38] V. Gaitsgory, "Control of Systems with Fast and Slow Motions", Nauka, Moscow, 1991 (in Russian).
- [39] V. Gaitsgory, "Suboptimization of Singularly Perturbed Control Problems", *SIAM J. Control and Optimization*, 30 (1992): 5, pp. 1228 - 1240.
- [40] V. Gaitsgory, "Limit Hamilton-Jacobi-Isaacs equations for singularly perturbed zero-sum differential games", *Journal of mathematical analysis and applications*, 202 (1996): 3, pp.862-899
- [41] V. Gaitsgory, "On Representation of the Limit Occupational Measures Set of a Control Systems with Applications to Singularly Perturbed Control Systems", *SIAM J. Control and Optimization*, 43 (2004): 1, pp 325-340
- [42] V. Gaitsgory and A. Leizarowitz, "Limit Occupational Measures Set for a Control System and Averaging of Singularly Perturbed Control Systems", *J. Math. Anal. and Appl.*, 233 (1999), pp. 461-475.
- [43] V. Gaitsgory and L. Manic, "Use of Approximations of Hamilton-Jacobi-Bellman Inequality for Solving Periodic Optimization Problems", arXiv:1309.1824 [math.OC] (<http://arxiv.org/abs/1309.1824>).
- [44] V. Gaitsgory and M.T. Nguyen, "Multiscale Singularly Perturbed Control Systems: Limit Occupational Measures Sets and Averaging", *SIAM J. Control and Optimization*, 41 (2002): 3, pp. 954-974.
- [45] V. Gaitsgory and M. Quincampoix, "Linear Programming Approach to Deterministic Infinite Horizon Optimal Control Problems with Discounting", *SIAM J. of Control and Optimization*, 48 (2009): 4, pp. 2480-2512.
- [46] V. Gaitsgory and M. Quincampoix, "On sets of occupational measures generated by a deterministic control system on an infinite time horizon", *Nonlinear Analysis Series A: Theory, Methods & Applications*, 88 (2013), pp 27-41.
- [47] V. Gaitsgory and S. Rossomakhine "Linear Programming Approach to Deterministic Long Run Average Problems of Optimal Control", *SIAM J. of Control and Optimization*, 44 (2005/2006): 6, pp. 2006-2007.
- [48] V. Gaitsgory and S. Rossomakhine "Averaging and Linear Programming in Some Singularly Perturbed Problems of Optimal Control", *Journal of Applied Mathematics and Optimization*, Published online in June 2014, DOI 10.1007/s00245-014-9257-1 (arXiv:1309.3734v2 [math.OC])
- [49] V. Gaitsgory, S. Rossomakhine and N. Thatcher, "Approximate Solutions of the HJB Inequality Related to the Infinite Horizon Optimal Control Problem with Discounting", *Dynamics of Continuous and Impulsive Systems Series B: Applications and Algorithms*, 19 (2012), pp. 65-92.
- [50] D. Goreac and O.S. Serea, "Linearization techniques for L^∞ - control problems and dynamic programming principles in classical and L^∞ control problems", *ESAIM: Control, Optimization and Calculus of Variations*, 18 (2012), pp. 836-859.
- [51] G. Grammel, "Averaging of Singularly Perturbed Systems", *Nonlinear Analysis*, 28 (1997), 1851-1865.

- [52] G. Grammel, "On Nonlinear Control Systems with Multiple Time Scales", *Journal of Dynamical and Control Systems*, 10 (2004), pp. 11-28.
- [53] D. Hernandez-Hernandez, O. Hernandez-Lerma, M. Taksar, "The Linear Programming Approach to Deterministic Optimal Control Problems", *Applicaciones Mathematicae*, 24 (1996): 1, pp 17-33.
- [54] O. Hernandez-Lerma and J.B. Lasserre "Markov Chains and Invariant Probabilities", Birkhauser-Verlag, Basel, 2003.
- [55] Y. Kabanov and S. Pergamenschikov, "Two-Scale Stochastic Systems", Springer Verlag, Berlin-Heidelberg, 2003.
- [56] P.V.Kokotovic, H.K.Khalil and J.O'Reilly, "Singular Perturbation Methods in Control: Analysis and Design", 2nd Edition, SIAM Classics in Applied Mathematics 25, 1999.
- [57] T.G. Kurtz and R.H. Stockbridge, "Existence of Markov Controls and Characterization of Optimal Markov Controls", *SIAM J. on Control and Optimization*, 36 (1998) :2, pp. 609-653.
- [58] H.J.Kushner, "Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems", Birkhauser, Boston, 1990.
- [59] H.J.Kushner and P.G. Dupuis, "Numerical Methods for Stochastic Control Problems in Continuous Time", 2nd (revised) edition, Springer-Verlag, New York, 2002.
- [60] J.B. Lasserre, D. Henrion, C. Prieur, and E. Trélat, "Nonlinear optimal control via occupation measures and LMI-relaxations", *SIAM J. Control Optim.*, 47 (2008), pp. 1643-1666.
- [61] A. Leizarowitz, "Order Reduction is Invalid for Singularly Perturbed Control Problems with Vector Fast Variables", *Math. Control Signals and Systems*, 15 (2002), pp. 101-119.
- [62] J.G. Llavona, "Approximation of Continuously Differentiable Functions", Mathematics Studies, 130, North Holland, Amsterdam, 1986.
- [63] S.D. Naidu, "Singular Perturbations and Time Scales in Control Theory and Applications: An Overview" , *Dynamics of Continuous Discrete and Impulsive Systems, Series B: Applications and Algorithms*, 9 (2002), pp. 233-278.
- [64] R.E. O'Malley, Jr. "Introduction to Singular Perturbations", Academic Press, New York, 1974.
- [65] R.E. O'Malley, Jr. "Singular Perturbations and Optimal Control", In Mathematical Control Theory, W.A. Copel, ed., Lecture Notes in Mathematics, 680, Springer-Verlag, Berlin 1978.
- [66] Pervozvanskii, A.A., Gaitsgori, V.G., 1988, "Theory of suboptimal decisions: Decomposition and aggregation", Kluwer, Dordrecht.
- [67] M. Quincampoix and F. Watbled, "Averaging method for discontinuous Mayer's problem of singularly perturbed control systems", *Nonlinear Analysis: Theory, Methods & Applications*, 54(2003), pp 819-837.
- [68] "Plotnikov V.A., Plotnikov A.V., Vituk A.N. Differential Equations with Multivalued Right-Hand Sides: Asymptotic Methods", AstroPrint, Odessa, 1999 (in Russian).
- [69] J.E. Rubio, "Control and Optimization. The Linear Treatment of Nonlinear Problems", Manchester University Press, Manchester, 1985.
- [70] O.S. Serea, "Characterization of the optimal trajectories for the averaged dynamics associated to singularly perturbed control systems", *Journal of Differential Equations*, 255(2013), pp. 42264243
- [71] R.H. Stockbridge, "Time-Average Control of a Martingale Problem. Existence of a Stationary Solution", *Annals of Probability*, 18 (1990), pp 190-205.
- [72] R.H. Stockbridge, "Time-Average Control of a Martingale Problem: A Linear Programming Formulation", *Annals of Probability*, 18 (1990), pp 206-217.
- [73] Vasil'eva A.B., Butuzov V.F., "Asymptotic Expansions of Solutions of Singularly Perturbed Equations", Nauka, Moscow, 1973 (in Russian).
- [74] V.Veliov, "A generalization of Tichonov Theorem for Singularly Perturbed Differential Inclusions", *Journal of Dynamical and Control Systems*, 3 (1997), pp 1-28
- [75] A.Vigodner, "Limits of Singularly Perturbed Control Problems: Dynamical Systems Approach, Thesis for the Degree of Doctor of Philosophy, The Weizmann Institute of Science, Rehovot, Israel, 1995.
- [76] R. Vinter, "Convex Duality and Nonlinear Optimal Control", *SIAM J. Control and Optimization* , 31 (1993): 2, pp. 518-538.
- [77] G.G. Yin and Q. Zhang, "Continuous-Time Markov Chains and Applications. A Singular Perturbation Approach", Springer, New York, 1997.